MITOCHONDRIAL AND RIBOSOMAL DNA ANALYSIS FOR IDENTIFICATION OF SIBLING SPECIES OF THE MOSQUITO, *Anopheles quadrimaculatus* (Say)

By

SHARON ELIZABETH MITCHELL

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1990

UNIVERSITY OF FLORIDA LIBRARIES

THIS WORK IS DEDICATED TO THE AUTHOR'S MOTHER,
ELIZABETH RUTH NEILL

ACKNOWLEDGEMENTS

The author wishes to express her sincere appreciation to Drs. J. A. Seawright and A. F. Cockburn for their critical appraisal of the dissertation, and for their guidance and friendship throughout the course of this study. Special thanks are extended to Dr. S. K. Narang for his aid in interpretation of results obtained from protein electrophoresis.

The author extends her gratitude to the graduate committee members Drs. J. E. Maruniak, H. G. Hall and C. D. Chase for their encouragement and critical review of the work presented. Very special thanks are extended to M. Q. Benedict for his help with the computer analyses and continued friendship throughout the years.

The author wishes to thank the staff of the U. S. Department of Agriculture, Insects Affecting Man and Animals Research Laboratory, Gainesville. She especially thanks Dr. G. A. Mount for his support of this study. Finally, the author extends special thanks to Ms. P. A. Martin for assistance in the preparation of this manuscript.

TABLE OF CONTENTS

PAGE

ACKNOWLEDGEMENTS iii

ABSTRACT v

CHAPTER I. INTRODUCTION 1

CHAPTER II. CLONING AND ORGANIZATION OF THE MITOCHONDRIAL GENOME OF *Anopheles quadrimaculatus* (Say) 4

Introduction 4
Materials and Methods 5
Results 8
Discussion 31

CHAPTER III. MITOCHONDRIAL AND RIBOSOMAL DNA VARIATION AMONG MEMBERS OF THE *Anopheles quadrimaculatus* SPECIES COMPLEX 36

Introduction 36
Materials and Methods 38
Results 43
Discussion 69

CHAPTER IV. CONCLUDING REMARKS 77

LITERATURE CITED 79

BIOGRAPHICAL SKETCH 84

Abstract of Dissertation Presented to the Graduate School of
the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

MITOCHONDRIAL AND RIBOSOMAL DNA ANALYSIS FOR IDENTIFICATION OF SIBLING SPECIES OF THE MOSQUITO, *Anopheles quadrimaculatus* (Say)

By

SHARON ELIZABETH MITCHELL

August 1990

Chairman: J. A. Seawright
Cochairman: A. F. Cockburn
Major Department: Entomology and Nematology

Work was done to investigate the feasibility of using DNA restriction patterns for the identification of individual mosquitoes of the *Anopheles quadrimaculatus* species complex. The entire 15 kilobase (kb) mitochondrial genome of *A. quadrimaculatus,* species A was cloned in bacteriophage, then subcloned in plasmid vectors and characterized. The cloned DNA was physically mapped with restriction endonucleases, and several genes were mapped by sequencing the ends of *A. quadrimaculatus* subclones and by hybridization with previously characterized *Aedes albopictus* mitochondrial DNA (mtDNA) clones. These portions of the genetic map were identical in gene order to those of *Drosophila yakuba.*

The cloned mtDNA and a previously isolated ribosomal DNA (rDNA) clone were used as hybridization probes for analysis of individual mosquitoes from seven populations containing more than one sibling species of *A. quadrimaculatus*. Hybridization of both the mtDNA and rDNA clones to total genomic DNA digested with several different restriction endonucleases produced species-specific restriction patterns and identified low-frequency intraspecific variation.

The mtDNA of the four sibling species of the *A. quadrimaculatus* complex were physically mapped with the same restriction enzymes employed in the population analysis. Homologies among the mtDNA restriction patterns confirmed the genetic relationships among the sibling species determined by isozyme analysis, chromosome comparisons and hybridization studies.

CHAPTER I

INTRODUCTION

Cryptic species occur commonly in the genus *Anopheles* (Diptera: Culicidae). Cryptic species are defined as genetically distinct, reproductively isolated but morphologically indistinguishable taxa within what was previously considered to be a single species (Narang and Seawright, 1990). The first species complex in anopheline mosquitoes was recognized about fifty years ago for *Anopheles maculipennis,* an important vector of human malaria in Europe (Kitzmiller, 1953). The discovery of this species complex resulted from the observation that even though *A. maculipennis* was widely distributed, the incidence of malaria was not. Further investigations indicated that there were ecological differences among different geographic strains of *A. maculipennis* and later these strains were found to correspond to eight sibling species based on egg morphology, hybridization studies and chromosomal comparisons (Kitzmiller et al., 1967). Since this initial discovery, species complexes have been described for numerous anophelines including *A. gambiae, A. culicifacies, A. balabacensis, A. punctulatus, A. annulipes, A. coustani, A. sinensis, A. maculatus, A. subpictus, A. marshallii* and *A. quadrimaculatus* (see Narang and Seawright, 1990 for a complete review). Some, but not all, members of these species complexes are important malaria vectors in Africa, India and Asia.

Although it is not common knowledge, especially in North America, the fact remains that malaria accounts for more morbidity and mortality, world-wide, than any other transmissible disease. The World Health Organization estimates that at any given time, about 100 million people show clinical symptoms of the disease and about 1 million die from malaria each year (Marshall, 1990).

From the epidemiological standpoint, it is important that only a few of the sibling forms in the typical anopheline species complex may be involved in the transmission of malaria parasites to humans. Therefore, correct identification of cryptic species is necessary so that control measures can be taken in areas where high densities of disease vectors occur and avoided if the nonvector, sibling species predominate.

Anopheles quadrimaculatus (Say), the common malaria mosquito, is widely distributed throughout the eastern and central parts of the United States and ranges north to southeastern Canada. When human malaria was prevalent in this country, *A. quadrimaculatus* was the primary vector of this disease in the southern and eastern-seaboard states. Presently, this mosquito is an important vector of canine heartworm, *Dirofilaria immitis,* throughout its range (Lewandowski et al., 1980).

Recently, it has been found that the taxon *A. quadrimaculatus* is a species complex containing at least four morphologically identical, sympatric sibling forms, species A, B, C, and D (Lanzaro et al., 1988; Kaiser et al., 1988a,b; Narang et al., 1989a,b). The genetic evidence for species designation of all four sibling types is conclusive and includes hybridization crosses, polytene chromosome analysis and diagnostic allozymes.

In addition to the traditional genetic techniques (i.e. cross-hybridization, chromosome analysis and isozyme electrophoresis), DNA restriction analysis has become a powerful method for studying the genetic structure of natural populations. The work reported herein applies the techniques of molecular biology to the practical problem of correct identification of morphologically indistinguishable species and to the more theoretical aspects of evolutionary genetics (i.e. genetic diversity and phylogenetic relationships among sibling species).

The objectives of this study were to (1) develop DNA probes for both cytoplasmic (mitochondrial) and nuclear (ribosomal) genomes of *A. quadrimaculatus* to be used as hybridization probes for the analysis of individual mosquitoes; (2) investigate the feasibility of using restriction site polymorphisms for both cytoplasmic and nuclear DNAs to identify individuals of the *A. quadrimaculatus* complex to species, and; (3) use the mitochondrial DNA restriction site data to estimate genetic distances and determine phylogenetic relationships among the four sibling species.

This dissertation contains four chapters, two of which are currently being submitted in modified form to scientific journals. Chapters II and III provide a short introduction, a review of the relevant literature, a presentation of the results of the work done and a discussion of the significance of the findings. A general summary and concluding remarks are contained in Chapter IV.

CHAPTER II

CLONING AND ORGANIZATION OF THE MITOCHONDRIAL GENOME OF *Anopheles quadrimaculatus*, SPECIES A

Introduction

Mitochondrial DNA (mtDNA) restriction analysis has become a powerful method for studying natural populations (Brown, 1985). In order to analyze the mtDNA of individual insects, it is useful to have purified or cloned mtDNA to use as a radioactive probe (Cockburn and Seawright, 1988). Isolation of mtDNA from mosquitoes in amounts sufficient for use as a probe is problematic because of the difficulty in obtaining large quantities of live insects and low yields. The complete mitochondrial genome has been cloned and sequenced from at least one insect, *Drosophila yakuba* (Clary and Wolstenholme, 1985). Although Dubin and coworkers have cloned portions of *Aedes albopictus* mtDNA (HsuChen et al., 1984; Dubin et al., 1986), cloning of the complete mitochondrial genome for mosquitoes has not been reported.

To facilitate restriction analysis of mtDNA polymorphisms in the *Anopheles quadrimaculatus* (Say) species complex, the entire mitochondrial genome from *A. quadrimaculatus*, species A was initially cloned as three Eco RI fragments in bacteriophage. A restriction map was prepared by digestion of the clones and confirmed by comparison to the digestion pattern of native mtDNA. Subsequently, mtDNA inserts from the bacteriophage clones were subcloned in plasmid vectors. Sequence data

obtained from 13 plasmid subclones allowed alignment of the physical map with the genetic map derived from Drosophila. The clones and physical map will allow comparison of mtDNA from the four species of the *A. quadrimaculatus* complex and determination of the relationship of their maternal lineages.

Materials and Methods

Adult *Anopheles quadrimaculatus,* sp. A (ORLANDO strain) were obtained from the USDA, Insects Affecting Man and Animals Research Laboratory, Gainesville, Florida. Total mosquito DNA was prepared as in Cockburn and Seawright (1988). Briefly, 2 grams of frozen mosquitoes were powdered in a mortar and pestle on dry ice and suspended in 100 ml of cold homogenization buffer (0.1 M NaCl, 0.2 M sucrose, 0.01 M ethylenediaminetetraacetic acid (EDTA), 0.03 M Tris HCl, pH 8.0). After homogenizing for a few strokes in a Wheaton dounce tissue grinder, the homogenate was transferred to a centrifuge bottle. Twenty-five ml of lysis buffer (0.25 M EDTA, 2.5% sodium dodecyl sulfate (SDS), 0.5 M Tris HCl, pH 9.2) and 1 mg proteinase K were added and the mixture was incubated at 55° C for 1 hr. Seventeen ml of 8 M KCH_3COOH was added, the bottle was placed on ice for 1 hr and then centrifuged at 17,000 g for 20 min. Three-hundred ml of ethanol was added to the supernatant and the solution was incubated at 0° C for 1 hr. Nucleic acids were pelleted at 17,000 g in a preparative centrifuge for 10 min and further purified by CsCl gradient centrifugation (Maniatis et al., 1987).

Mitochondria were prepared by differential centrifugation. About 2 g of live mosquitoes were thoroughly homogenized (Wheaton dounce homogenizer) in 100 ml

homogenization buffer on ice and centrifuged at 1,600 g for 5 min. The supernatant was decanted and nuclei pelleted by centrifugation at 1,600 g for 10 min. Mitochondria were pelleted by centrifuging the supernatant at 17,000 g for 30 min. The mitochondrial pellet was resuspended in 25 ml homogenization buffer, centrifuged at 1,600 g for 10 min, the supernatant decanted, and centrifuged at 40,000 g for 10 min. The mitochondrial pellet was suspended in 8 ml of homogenization buffer and lysed by the addition of 0.5 ml 20% sodium lauryl sarcosinate. Cesium chloride (CsCl) (8.5 g) and 0.4 ml 10 mg/ml ethidium bromide were added, the solution was transferred to an ultracentrifuge tube and centrifuged in a Beckman L8-M ultracentrifuge with a type 70.1 Ti fixed rotor at 35,000 RPM for 40 hr. The lower mtDNA band was visualized by UV irradiation and removed. Ethidium bromide was extracted and purified mtDNA precipitated using standard techniques (Maniatis et al., 1987).

Purified Eco RI digested mtDNA was cloned into the Eco RI site of bacteriophage λ gt10 (Stratagene) using standard procedures (Berger and Kimmel, 1987). DNA was packaged into bacteriophage according to the manufacturer's instructions using Gigapack Gold packaging mix (Stratagene). Agarose gel electrophoresis, nitrocellulose plaque lifts, Southern transfer and nick translation of DNA were standard techniques (Berger and Kimmel, 1987). Hybridizations were as in Cockburn and Seawright (1988).

Isolation of bacteriophage DNA, subcloning of fragments, transformation, plasmid growth, isolation of plasmid DNA, and restriction digestion were standard techniques (Berger and Kimmel, 1987). Plasmids were grown in *Escherichia coli* strain DH5α (BRL). Mitochondrial DNA inserts from bacteriophage clones were purified from

agarose gels using the GENECLEAN™ system according to the manufacturer's instructions (BIO 101).

Two non-overlapping *Aedes albopictus* mtDNA clones, k8 and k14, were obtained from Dr. Donald T. Dubin (University of Medicine and Dentistry of New Jersey).

Plasmid subclones were sequenced by the dideoxy chain-termination method using the Sequenase version 2.0 kit from United States Biochemical Corporation. The ends of all subclones were sequenced at least twice except when sequence data had already been obtained from another clone which shared a common end. Sequencing was done in one direction only. Therefore, sequence data was obtained for both mtDNA strands only for the middle sections of relatively small subclones (i.e. clones p3X2 and p3RX2).

Alignment of *A. quadrimaculatus* and *D. yakuba* mitochondrial sequences was done with the aid of the Genetics Computer Group Sequence Analysis Software Package (Devereaux et al., 1984). The *D. yakuba* sequence (Genbank) was searched and regions of sequence similarity were identified using the WORDSEARCH program (Wilbur and Lipman, 1983). Final alignment and calculation of percent similarity between the two sequences was accomplished with the BESTFIT sequence comparison program (Smith and Waterman, 1981). The MAP program was used to translate both mosquito and *Drosophila* sequence data for protein comparisons. The *Drosophila* mtDNA code was used to predict protein sequences. The validity of sequence alignments in the A+T-rich region was evaluated by randomizing this section of DNA from *D. yakuba* using the SHUFFLE program. Twenty randomized sequences were generated and each of these was realigned with the mosquito sequence.

Results

Cloning

The complete mitochondrial genome of animals is frequently difficult to clone in bacteria. For example, an *A. quadrimaculatus* total genomic DNA library was previously screened for mtDNA clones. The library was made by cloning total DNA that had been partially digested with *Sau* 3A into the bacteriophage λ vector EMBL 3A. Although the mtDNA represents about 1% of the total DNA, contains *several Sau* 3A restriction sites, and would be near the optimal size to insert into this vector, no mtDNA inserts were found in about 50,000 plaques screened. About 500 mitochondrial clones should have been found if there was no bias against them.

Other workers have noted that cloning of mammalian mtDNA that is unstable in bacteria can be successful if the DNA is cloned as relatively small fragments (W. Hauswirth, personal communication). Previous experiments involving hybridization of *Eco* RI digested total genomic DNA from *Anopheles quadrimaculatus* sp. A to purified mtDNA revealed that the mtDNA of this species contains three *Eco* RI fragments (6.3, 5.6 and 3.4 kb). Therefore, an *Eco* RI library of purified mtDNA was constructed in λ gt10. This library was screened with two non-overlapping *Aedes albopictus* mtDNA clones, k8 and k14. Two clones were isolated, one (AqAm1) homologous to k14 and one (AqAm2) homologous to k8. The *Eco* RI inserts in these were subcloned into the plasmid vector pIBI24. The mitochondrial library was rescreened with total mtDNA from *A. quadrimaculatus* and counter-screened with a mixture of the plasmid subclones (pAYCm1 and pAYCm2). A clone (AqAm3) was isolated that hybridized to mtDNA but not the other two clones. This clone was subcloned into pIBI24 to form pAYCm3.

The three lambda clones each contained one of the three *Eco* RI fragments and apparently represented the entire mitochondrial genome.

Analysis of the inserts from the three bacteriophage clones indicated that these *Eco* RI fragments were identical in size to the native mtDNA fragments (data not shown). Inserts from two of the three plasmid subclones, however, were not identical to the parent clones (Figure 2-1). One subclone (pAYCm1) contained a 5.1 kb insert (about 1.2 kb had been deleted from one end). The deletion from the end of pAYCm1 was probably due to cleavage at an internal *Eco* RI* site, since an Eco RI site was regenerated at this end. Another clone (pAYCm3) had a 6.3 kb insert (an extra 0.7 kb was added to the interior of the clone). The 3.4 kb insert of the third subclone (pAYCm2) was unaltered.

Inserts of AqAm1 and AqAm3 were cleaved into smaller fragments with several different enzymes and, again, subcloned into plasmids. Because of poor growth and low DNA yields of the pIBI-derived plasmids, pUC 19 was used as the plasmid vector. As a result of this approach, several subclones were isolated from each bacteriophage (see Figure 2-2), comprising the complete insert of each. The mtDNA insert from pAYCm2 was transferred into pUC 19, as well, to form p2R1.

The initial attempts to subclone the right end of AqAm3 failed. Several hundred lac$^-$ recombinant subclones from three separate digests (*Eco* RI/*Stu* I, *Eco* RI/*Sst* I and *Eco* RI/*Xba* I) were isolated by X-gal selection. The identity of the subcloned fragment was determined by digestion with the appropriate restriction endonuclease and gel electrophoresis. The left end of the mtDNA and various vector fragments were recovered many times from each digest, but the mtDNA sequences immediately adjacent to the righthand *Eco* RI site were never obtained. Fortuitously, one subclone having a

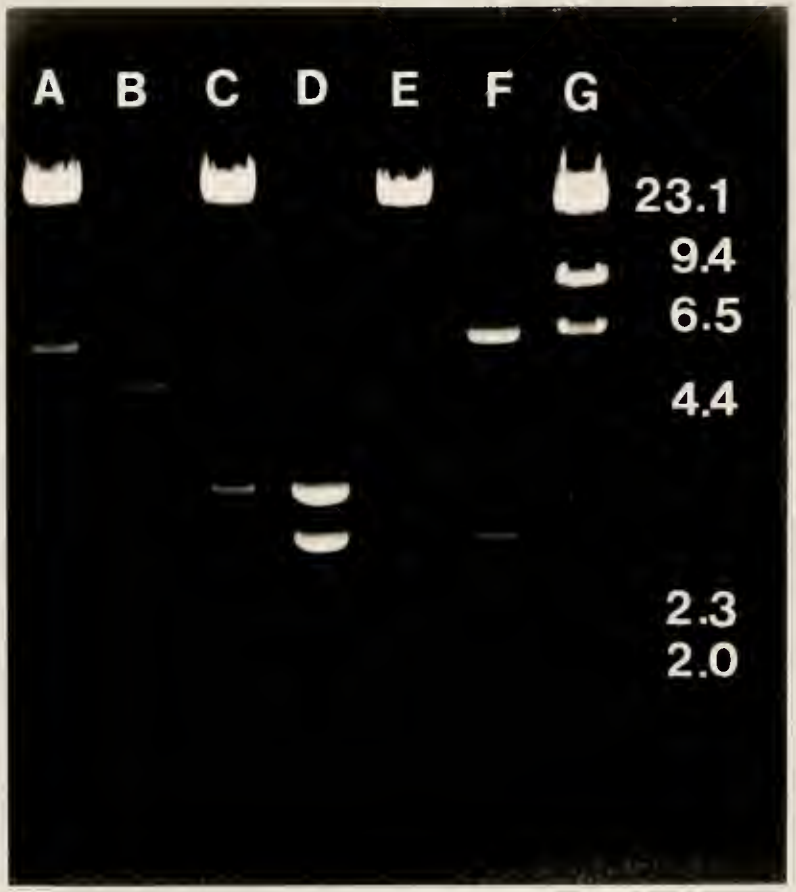


Figure 1-1. Eco R1 digests of λgt 10 and pIBI-24 mtDNA clones from *Anopheles quadrimaculatus.*

Lane (A) AqAm1; (B) pAYCm1; (C) AqAm2; (D) pAYCm2; (E) AqAm3; (F) pAYCm3; (G) λ Hind III fragments. Note that the 32.7 kb and 10.6 kb fragments in lanes A, C and E are λ vector arms. Also the 2.9 kb fragment in lanes B, D and F are the pIBI-24 vector.

partially nonrecombinant phenotype (lac$^+$) was selected for analysis. This clone was from an *Eco* RI/*Xba* I digest and contained the rightmost *Eco* RI/*Xba* I mtDNA fragment (0.6 kb) plus the noncontiguous 2.3 kb *Xba* I mtDNA fragment. This subclone was digested with *Xba* I and religated to make subclone p3RX2 (Figure 2-2), which was strongly lac$^+$. Further screening of lac$^+$ subclones yielded a subclone containing the 1.9 kb *Eco* RI/*Stu* I mtDNA fragment (p3RU2). Subclones containing the 4.2 kb *Eco* RI/*Sst* I fragment (also lac$^+$) were rearranged. All of these lac$^+$ subclones had the same *Eco* RI end (internal to the large ribosomal RNA coding sequence) ligated to the plasmid β-galactosidase α-peptide, so presumably the mitochondrial sequences upstream from that site are responsible for the restoration of the lac$^+$ phenotype.

Restriction Analysis

Because of the initial difficulty encountered in obtaining unrearranged plasmid clones, restriction sites were mapped in the three bacteriophage clones. Vector DNA and DNA from the clones were cleaved with the appropriate restriction endonuclease(s), loaded in adjacent wells and electrophoresed in 0.8% agarose gels. Single and double digests were done for a total of 12 restriction enzymes. Ambiguities were clarified by repeating the appropriate digests with purified mtDNA inserts excised from agarose gels. Clones were oriented and the resulting composite map was verified by comparison to restriction digests of total cellular DNA blotted to nitrocellulose and probed with mtDNA. For example, pAYCm1 contained two *Eco* RI/*Ava* I fragments (5.2 and 1.1 kb). Clone pAYCm2 also contained two such fragments (2.2 and 1.2kb). To comprise the 3.3 kb *Ava* I fragment observed in native mtDNA (see Figure 2-3, third lane), the

1.1 kb *Eco* RI/*Ava* I of pAYCm1 and the 2.2 kb fragment of pAYCm2 must be adjacent to one another.

This restriction mapping indicated that the three mtDNA clones had the same profile as total cellular mtDNA (Figure 2-2). The integrity of all plasmid subclones subsequently isolated was evaluated with respect to the restriction map obtained from the bacteriophage clones and by DNA sequencing.

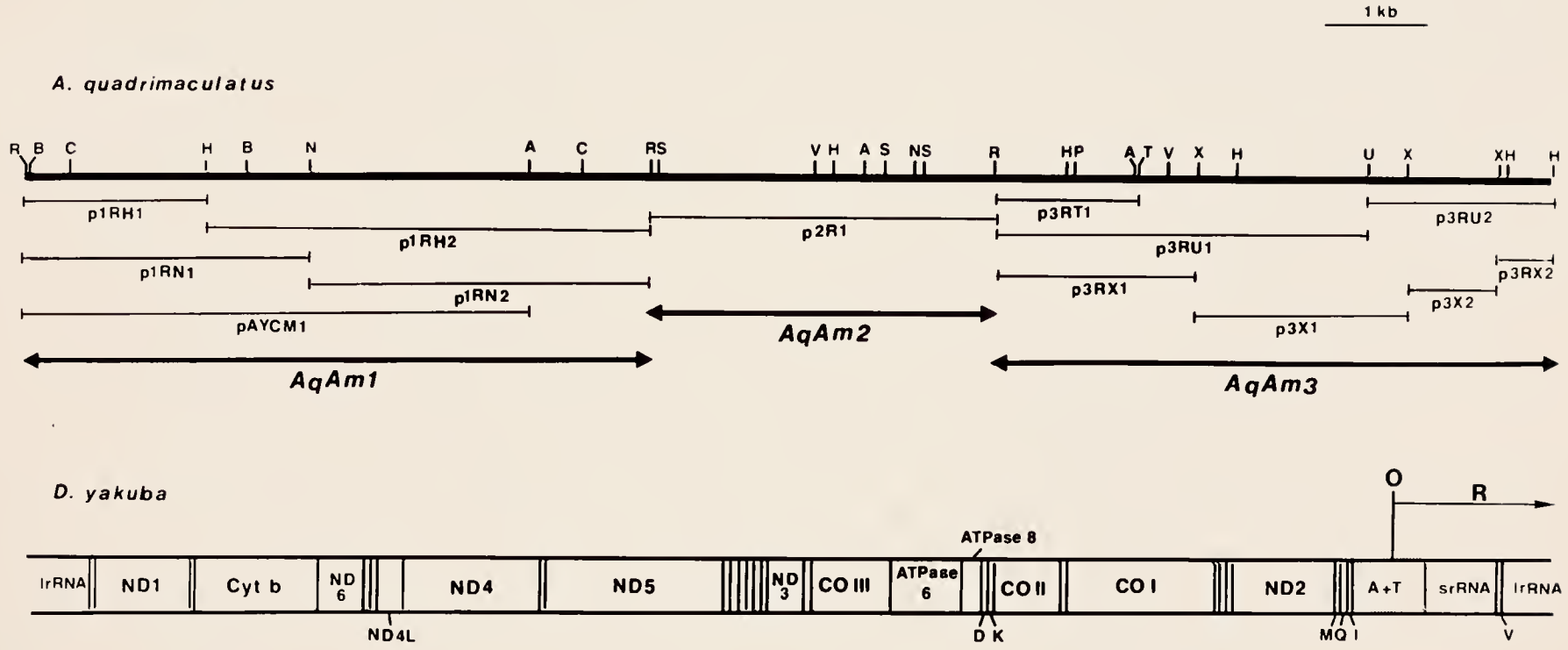
Mapping of *Aedes albopictus* mtDNA clones

Animal mitochondrial DNAs studied so far code for 2 ribosomal RNAs (rRNAs), 22 tRNAs, an A+T-rich region which has no corresponding transcript(s) and contains the origin of mtDNA replication and 13 hydrophobic proteins (Garesse, 1988). The protein coding regions code for subunits of enzyme complexes associated with the inner mitochondrial membrane and include cytochrome b (Cytb), two subunits of the ATPase (6 and 8), three subunits of the cytochrome c oxidase (CO I, II and III) and seven open reading frames which have recently been identified as subunits of the NADH reductase complex (ND1, 2, 3, 4, 4L, 5 and 6) (Chomyn et al., 1985).

Dubin and coworkers have previously cloned and partially sequenced three small parts of the *A. albopictus* mtDNA. These consisted of 1) 2.8 kb containing the 3' end of the *CO III* gene, $tRNA^{Gly}$, the *ND 3* gene, a cluster of tRNAs ($tRNA^{Arg}$, $tRNA^{Ala}$, $tRNA^{Asn}$, $tRNA^{Ser}$, $tRNA^{Glu}$ and $tRNA^{Phe}$) and the 3' end of the *ND 5* gene (clone k8) (Dubin et al., 1986); 2) the *Hind* III D fragment, 1.9 kb containing the 3' end of the *large subunit rRNA* gene (*lrRNA*), $tRNA^{Leu}$, and the 5' end of the *ND 1* gene (clone k14)

Figure 2-2. Restriction map of the *Anopheles quadrimaculatus*, species A mitochondrial genome.

The mitochondrial map is circular and has been opened at a selected *Eco* RI site. Positions of plasmid subclones and λ clones are noted beneath the restriction map. The genetic map from *Drosophila yakuba* is presented at the bottom of the page. Note that the *Drosophila* A+T-rich region has been shorted 0.7 kb to allow alignment with the mosquito genome. A=*Ava* I; B=*Bgl* II; C=*Cla* I; H=*Hind* III; N=*Nsi* I; P=*Pst* I; R=*Eco* RI; S=*Sca* I; T=*Sst* I; U=*Stu* I; V=*Pvu* II; X=*Xba* I.


1 kb
A. quadrimaculatus
R B C H B N A C RS V H A S NS R HP A T V X H U X XH H
p1RH1
p1RH2
p2R1
p3RT1
p3RU2
p3RU1
p1RN1
p1RN2
pAYCM1
p3RX1
p3RX2
p3X2
p3X1
AqAm1
AqAm2
AqAm3
D. yakuba
O
R
ATPase 8
lrRNA
ND1
Cyt b
ND 6
ND4
ND5
ND 3
CO III
ATPase 6
CO II
CO I
ND2
A+T
srRNA
lrRNA
ND4L
D K
MQ I
V

(HsuChen et al., 1984); and 3) the *Hind* III E fragment, 0.5 kb containing the 5' end of the *large ribosomal RNA (lrRNA)* gene and $tRNA^{Val}$ (HsuChen et al. 1984).

Two of these cloned DNA sequences (k8 and k14) were hybridized to total *A. quadrimaculatus* DNA cleaved with various enzymes to determine the regions of homology. Plasmid subclone k8 hybridized to fragments contained in AqAm2, including the 3.4 kb *Eco* RI fragment, the 6.3 kb and 2.2 kb *Hind* III fragments, the 3.3 kb *Ava* I fragment, the 10.2 kb *Cla* I fragment, the 13.0 kb *Bgl* II fragment, and the 13.1 kb and 2.2 kb *Sca* I fragments (Figure 2-3). Therefore, the k8 homologous region extends from just left of the *Eco* RI site between the 3.4 kb and 6.3 *Eco* RI fragments to the *Ava* I site between the 3.3 kb and 2.5 kb *Ava* I fragments. These results are consistent with the restriction map generated from the cloned *A. quadrimaculatus* mtDNA and with the isolation of AqAm2 using a k8 probe.

Clone k14 hybridized to the 6.3 kb *Eco* RI fragment, the 1.85 kb *Hind* III fragment, the 9.2 kb *Ava* I fragment, the 10.2 kb and 5.1 kb *Cla* I fragments, the 2.3 kb and 13.0 kb *Bgl* II fragments, and the 13.1 kb *Sca* I fragment (Figure 2-3). The k14 homologous region appears to be identical to the 1.85 kb *Hind* III fragment, which is found in all four species of the *A. quadrimaculatus* complex. In fact, the restriction patterns of mtDNAs from the two mosquito genera are very similar in the region of the *lrRNA* gene. The three *Hind* III sites that create the D and E fragments appear to be conserved as well as the *Eco* RI site in the *Hind* III D fragment.

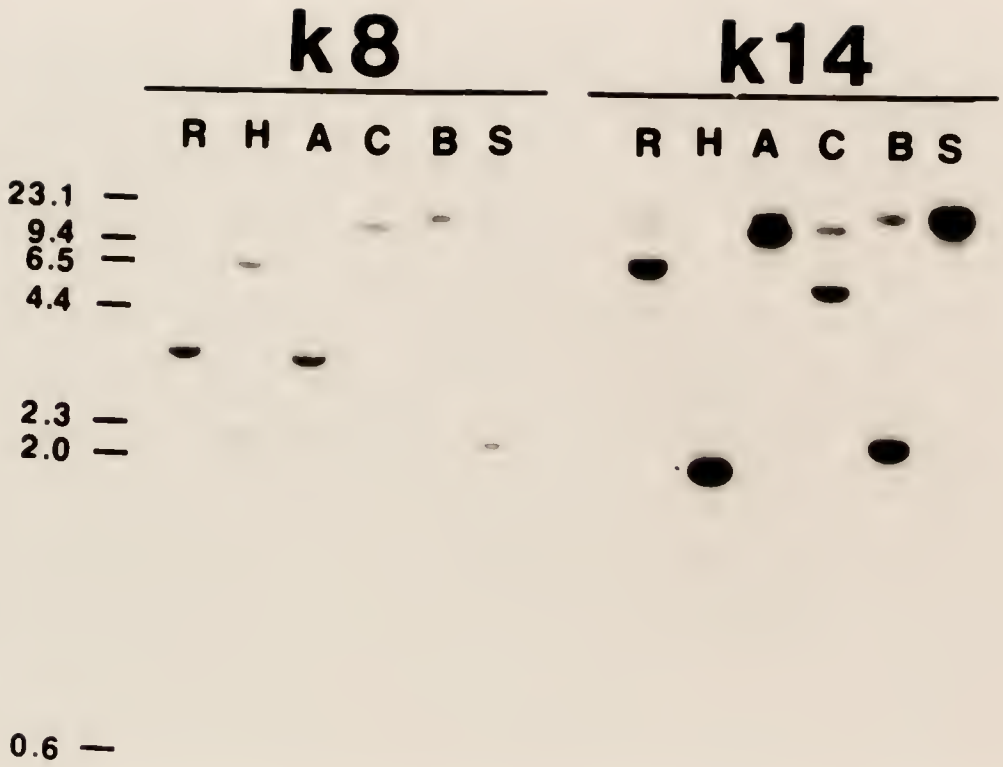


Figure 2-3. Total genomic DNA from *Anopheles quadrimaculatus* digested with various restriction enzymes and probed separately with the k8 and k14 mtDNA clones from *Aedes albopictus.*

Note: Abbreviations for the restriction enzymes are the same as in Figure 2-2. Positions of the λ *Hind* III marker fragments are noted in the far left column.

Sequencing of the ends of subcloned *A. quadrimaculatus* mtDNA

Sequence information was obtained from the ends of 13 plasmid subclones as well as an internal region of AqAm1 (the right end of pAYCm1) (Table 2-1; Figures 2-4 and 2-5). The sequencing results and the cross-hybridization with the *A. albopictus* clones allowed the orientation of the physical restriction map with the *D. yakuba* genetic map derived from the complete sequence information (Clary and Wolstenholme, 1985).

Seven out of 13 protein coding sequences, both rRNA coding sequences, 6 out of 22 tRNA coding sequences, and the A+T rich region were localized and occupied the same positions as in *Drosophila* (Figure 2-2). When these results are combined with the information derived from the *A. albopictus* clones, it appears that the gene order in *Drosophila* and mosquitoes is identical.

In general, there was considerable homology at the DNA level between the mosquito and *Drosophila* sequences in coding regions (Table 2-1). Similarity at the DNA level ranged from 86.8 % for the 5' end of the *CO I* gene to 71.4 % for the 3' end of the *ND 5* gene. Significant homology is also seen in the predicted amino acid sequences of the protein coding regions. It should be noted that percent similarity on the protein level (Table 2-1) was calculated on an absolute basis, without correcting for conservative substitutions. Therefore, this similarity figure substantially underestimates the structural similarity of these proteins. The major difference observed in protein coding sequences of the two organisms is that the *A. quadrimaculatus ND 1* gene apparently terminates 11 amino acids upstream of the *D. yakuba ND 1* gene (Figure 2-4l). It is not known if this discrepancy represents a real difference in the protein product, a cloning artifact, or the use of alternative stop codons in the mosquito.

TABLE 2-1. Sequence analysis summary of *Anopheles quadrimaculatus* mtDNA subclones.

Plasmid subclone	Insert size	ends	Sequenced bp	end	*D. yakuba* Alignment		Gene(s)	% Similarity DNA	Protein
p3RU1	3.7kb	R/U[a]	199	U	1	239	I,Q,M	92.5	-
p3X1	2.3kb	X/X	303	X^L	1,481	1,783	CoI	86.8	97.0
p3RX1	1.9kb	R/X	422	X	1,807	2,228	CoI	84.6	96.4
p3RT1	1.4kb	R/T	342	T	2,339	2,680	CoI	86.5	93.8
p3RU1	3.7kb	R/U	349	R	3,338	3,686	CoII	81.4	83.6
p2R1	3.4kb	R/R	212	R^R	3,738	3,943	CoII,K,D,Atp8[d]	85.4	-
p2R1	3.4kb	R/R	283	R^L	6,874	7,156	ND5	71.4	72.3
p1RN2	3.4kb	R/N	306	R	7,232	7,537	ND5	83.0	90.1
pAYCm1	5.2kb	R/R*[b]	236	R*	8,294	8,531	ND4	81.5	79.7
p1RN1	2.9kb	R/N	310	N	10,592	10,901	Cytb	77.7	73.3
p1RH2	4.45kb	R/H	292	H	11,277	11,568	Cytb	84.9	83.5
p1RH1	1.85kb	R/H	295	H	11,696	11,985	ND1	81.0	74.4
p1RN1	2.9kb	R/N	215	R	13,178	13,394	lrRNA	87.4	-
p3RX2	0.7kb	R/X	668	R/X	13,478	14,142	lrRNA	81.5	-
p3X2	0.8kb	X/X	771	X/X	14,177	14,939	srRNA	81.0	-
p3RU1	3.7kb	R/U	189	U	-[c]	-	A+T	-	-
p3X1	2.3kb	X/X	192	X	-	-	A+T	-	-

[a] Letter designations for restriction sites are as in legend for Figure 1-1.
[b] R* refers to the presumptive EcoRI* site.
[c] Alignments in A+T-rich region were not significant.
[d] Insufficient protein sequence was obtained from the ends of the CoII and Atp8 genes to determine protein similarity.
[L] Left end.
[R] Right end.

Figure 2-4. Mitochondrial DNA sequence alignments and predicted protein sequence for the *A. quadrimaculatus* clones.

(a) Alignment of p3RU1 (right end, *Stu* I); (b) Alignment and predicted amino acid sequence of p3X1 (left end, Xba I); (c) Alignment and predicted amino acid sequence of p3RX1 (right end, *Xba* I); (d) Alignment and predicted amino acid sequence of p3RT1 (right end, *Sst* I); (e) Alignment and predicted amino acid sequence of p3RU1 (left end, *Eco* RI); (f) Alignment and predicted amino acid sequence of p2R1 (right end, *Eco* RI); (g) Alignment and predicted amino acid sequence of p2R1 (left end, *Eco* RI); (h) Alignment and predicted amino acid sequence of p1RN2 (Right end, *Eco* RI); (i) Alignment and predicted amino acid sequence of pAYCm1 (left end, *Eco* RI*); (j) Alignment and predicted amino acid sequence of p1RN1 (right end, *Nsi* I); (k) Alignment and predicted amino acid sequence of p1RH2 (left end, *Hind* III); (l) Alignment and predicted amino acid sequence of p1RH1 (right end, *Hind* III); (m) Alignment of p1RN1 (right end, *Eco* RI); (n) Alignment of p3RX1 (entire clone); (o) Alignment of p3X2 (entire clone).

Note: The mtDNA sequence of *D. yakuba* appears beneath the mosquito sequence. Only differences in base composition are noted for the *Drosophila* sequence. Dots represent sequence gaps and dashes denote sequence identity. Direction of transcription is indicated by an arrow. Boxes indicate the boundaries of the *Drosophila* tRNA genes. The anticodons of tRNAs are underlined. Amino acids are designated by the standard, one-letter code. A=Alanine, R=Arginine, N=Asparagine, D=Aspartic acid, C=Cysteine, E=Glutamic acid, Q=Glutamine, G=Glycine, H=Histidine, I=Isoleucine, L=Leucine, K=Lysine, M=Methionine, F=Phenylalanine, P=Proline, S=Serine, T=Threonine, W=Tryptophan, Y=Tyrosine, V=Valine, *=Stop.

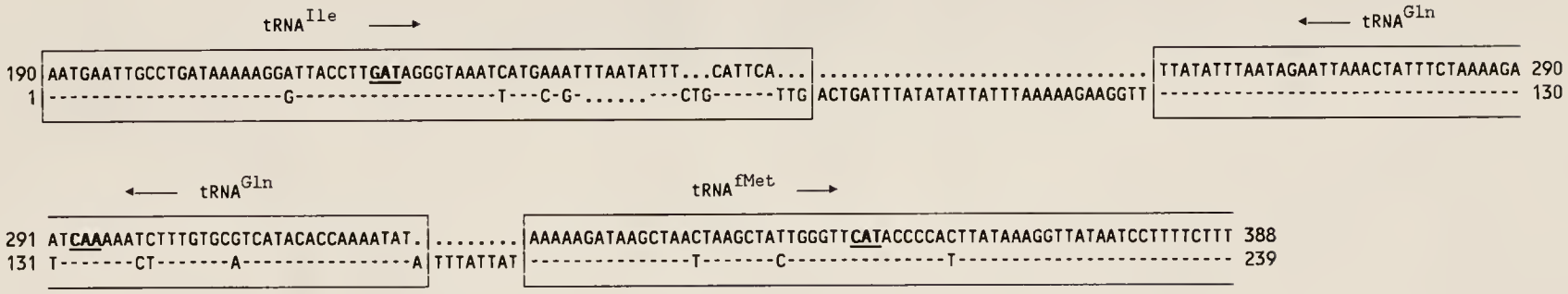


Figure 1-4a. p3RU1 (right end, Stu I)

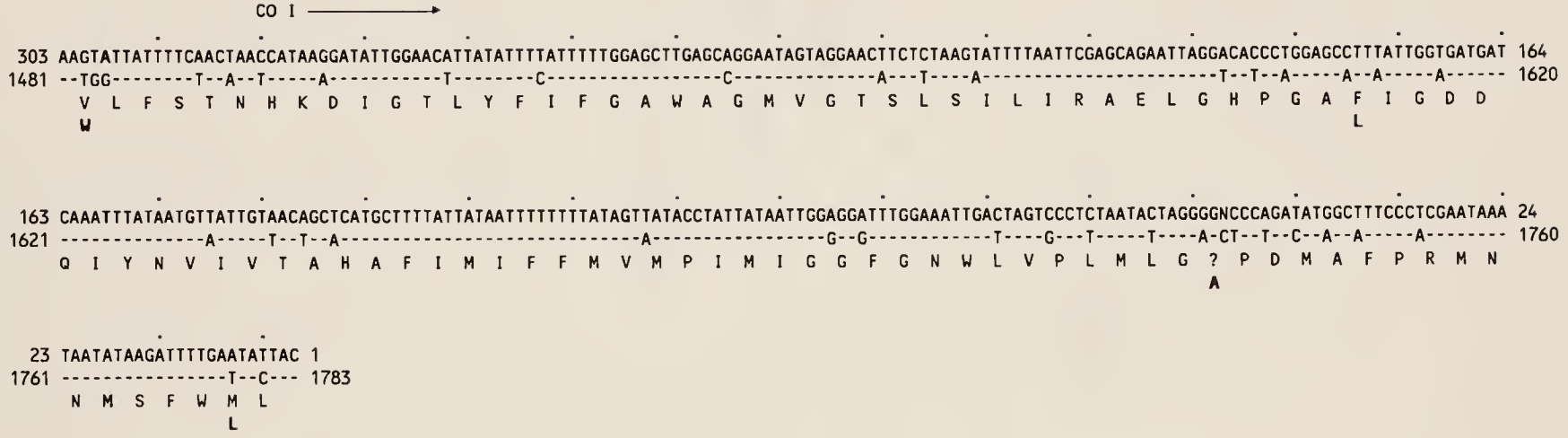

CO I
303 AAGTATTATTTTCAACTAACCATAAGGATATTGGAACATTATATTTTATTTTTGGAGCTTGAGCAGGAATAGTAGGAACTTCTCTAAGTATTTTAATTCGAGCAGAATTAGGACACCCTGGAGCCTTTATTGGTGATGAT 164
1481 --TGG--------T--A--T-----A----------T--------C----------C----------A---T----A----------T--T--A-----A--A-----A------ 1620
V L F S T N H K D I G T L Y F I F G A W A G M V G T S L S I L I R A E L G H P G A F I G D D
W
L
163 CAAATTTATAATGTTATTGTAACAGCTCATGCTTTTATTATAATTTTTTTTTATAGTTATACCTATTATAATTGGAGGATTTGGAAATTGACTAGTCCCTCTAATACTAGGGGNCCCAGATATGGCTTTCCCTCGAATAAA 24
1621 ----------A-----T--T--A----------A----------G--G----------T----G---T-----T----A-CT--T--C--A--A-----A-------- 1760
Q I Y N V I V T A H A F I M I F F M V M P I M I G G F G N W L V P L M L G ? P D M A F P R M N
A
23 TAATATAAGATTTTGAATATTAC 1
1761 ----------T--C--- 1783
N M S F W M L
L

Figure 1-4b. p3X1 (left end, Xba I).

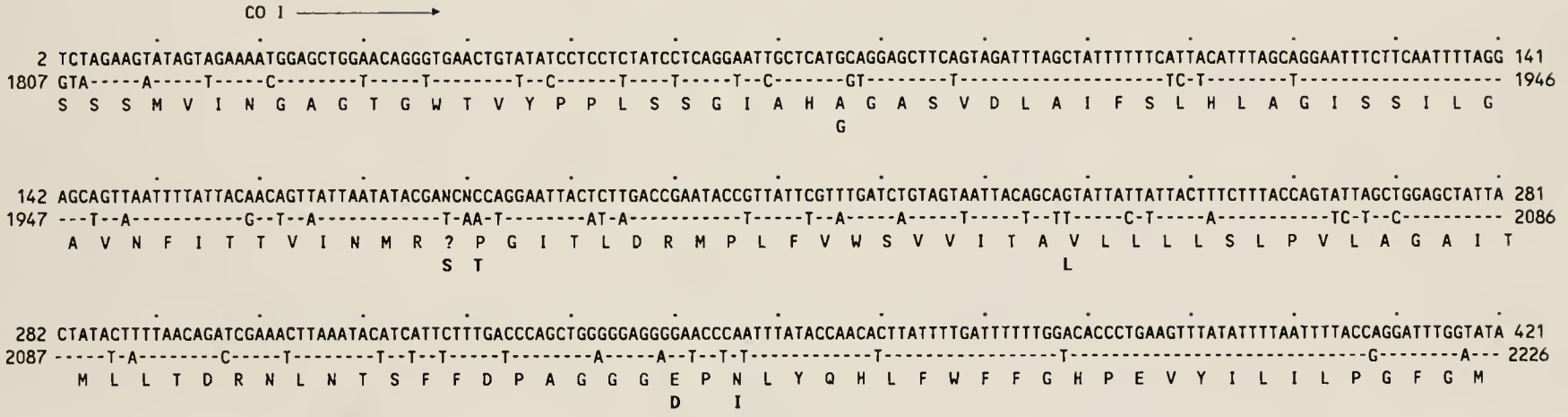

CO I
2 TCTAGAAGTATAGTAGAAAATGGAGCTGGAACAGGGTGAACTGTATATCCTCCTCTATCCTCAGGAATTGCTCATGCAGGAGCTTCAGTAGATTTAGCTATTTTTTCATTACATTTAGCAGGAATTTCTTCAATTTTAGG 141
1807 GTA-----A-----T-----C--------T-----T--------T--C------T----T-----T--C-------GT--------T--------------------TC-T--------T-------------------- 1946
S S S M V I N G A G T G W T V Y P P L S S G I A H A G A S V D L A I F S L H L A G I S S I L G
G
142 AGCAGTTAATTTTATTACAACAGTTATTAATATACGANCNCCAGGAATTACTCTTGACCGAATACCGTTATTCGTTTGATCTGTAGTAATTACAGCAGTATTATTATTACTTTCTTTACCAGTATTAGCTGGAGCTATTA 281
1947 ---T--A----------G--T--A-----------T-AA-T--------AT-A----------T-----T--A-----A-----T-----T--TT-----C-T-----A----------TC-T--C---------- 2086
A V N F I T T V I N M R ? P G I T L D R M P L F V W S V V I T A V L L L L S L P V L A G A I T
S T L
282 CTATACTTTTAACAGATCGAAACTTAAATACATCATTCTTTGACCCAGCTGGGGGAGGGGAACCCAATTTATACCAACACTTATTTTGATTTTTTGGACACCCTGAAGTTTATATTTTAATTTTACCAGGATTTGGTATA 421
2087 -----T-A--------C-----T--------T--T--T-----T--------A-----A--T--T-T----------T----------T----------G--------A--- 2226
M L L T D R N L N T S F F D P A G G G E P N L Y Q H L F W F F G H P E V Y I L I L P G F G M
D I

Figure 1-4c. p3RX1 (right end, Xba I).

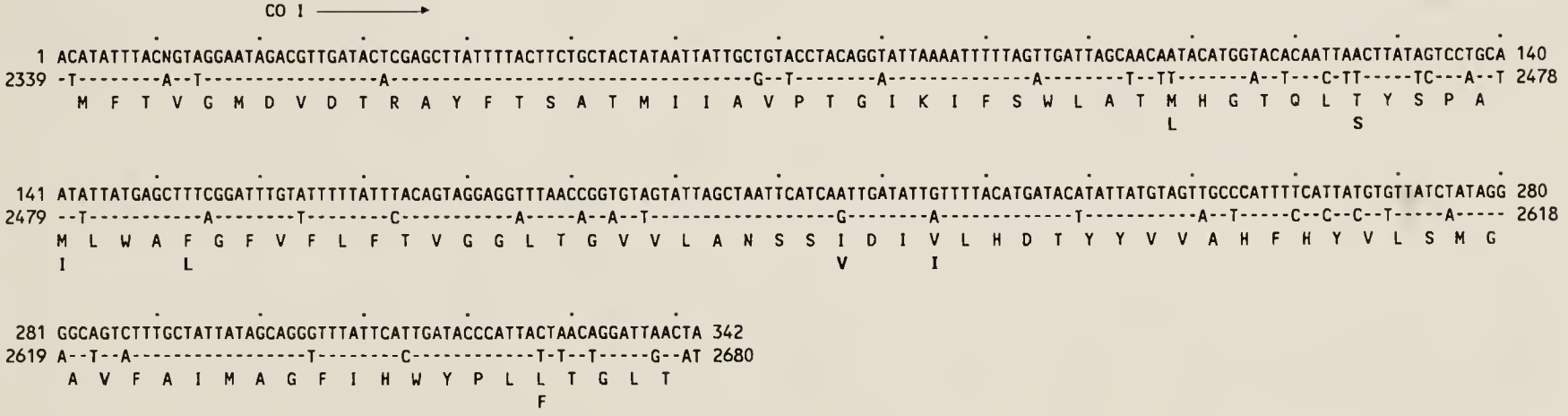

CO I
1 ACATATTTACNGTAGGAATAGACGTTGATACTCGAGCTTATTTTACTTCTGCTACTATAATTATTGCTGTACCTACAGGTATTAAAATTTTTAGTTGATTAGCAACAATACATGGTACACAATTAACTTATAGTCCTGCA 140
2339 -T--------A--T----------------A----------G--T--------A----------A--------T--TT-------A--T---C-TT-----TC---A--T 2478
M F T V G M D V D T R A Y F T S A T M I I A V P T G I K I F S W L A T M H G T Q L T Y S P A
L S
141 ATATTATGAGCTTTCGGATTTGTATTTTTATTTACAGTAGGAGGTTTAACCGGTGTAGTATTAGCTAATTCATCAATTGATATTGTTTTACATGATACATATTATGTAGTTGCCCATTTTCATTATGTGTTATCTATAGG 280
2479 --T----------A--------T--------C----------A-----A--A--T----------G--------A----------T----------A--T-----C--C--C--T-----A----- 2618
M L W A F G F V F L F T V G G L T G V V L A N S S I D I V L H D T Y Y V V A H F H Y V L S M G
I L V I
281 GGCAGTCTTTGCTATTATAGCAGGGTTTATTCATTGATACCCATTACTAACAGGATTAACTA 342
2619 A--T--A----------T--------C----------T-T--T-----G--AT 2680
A V F A I M A G F I H W Y P L L T G L T
F

Figure 1-4d. p3RT1 (right end, Sst I).

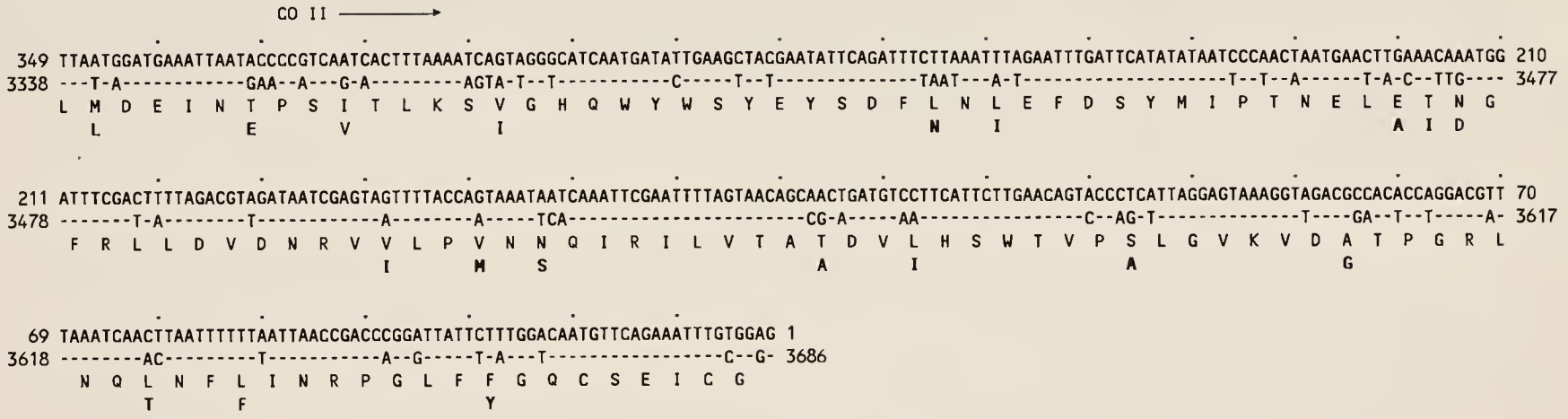

CO II
349 TTAATGGATGAAATTAATACCCCGTCAATCACTTTAAAATCAGTAGGGCATCAATGATATTGAAGCTACGAATATTCAGATTTCTTAAATTTAGAATTTGATTCATATATAATCCCAACTAATGAACTTGAAACAAATGG 210
3338 ---T-A----------GAA--A---G-A---------AGTA-T--T----------C-----T--T----------TAAT---A-T----------T--T--A------T-A-C--TTG---- 3477
L M D E I N T P S I T L K S V G H Q W Y W S Y E Y S D F L N L E F D S Y M I P T N E L E T N G
L E V I N I A I D
211 ATTTCGACTTTTAGACGTAGATAATCGAGTAGTTTTACCAGTAAATAATCAAATTCGAATTTTAGTAACAGCAACTGATGTCCTTCATTCTTGAACAGTACCCTCATTAGGAGTAAAGGTAGACGCCACACCAGGACGTT 70
3478 -------T-A--------T----------A--------A-----TCA----------CG-A-----AA----------C--AG-T----------T----GA--T--T-----A- 3617
F R L L D V D N R V V L P V N N Q I R I L V T A T D V L H S W T V P S L G V K V D A T P G R L
I M S A I A G
69 TAAATCAACTTAATTTTTTAATTAACCGACCCGGATTATTCTTTGGACAATGTTCAGAAATTTGTGGAG 1
3618 --------AC---------T----------A--G-----T-A---T----------C--G- 3686
N Q L N F L I N R P G L F F G Q C S E I C G
T F Y

Figure 1-4e. p3RU1 (left end, Eco R1).

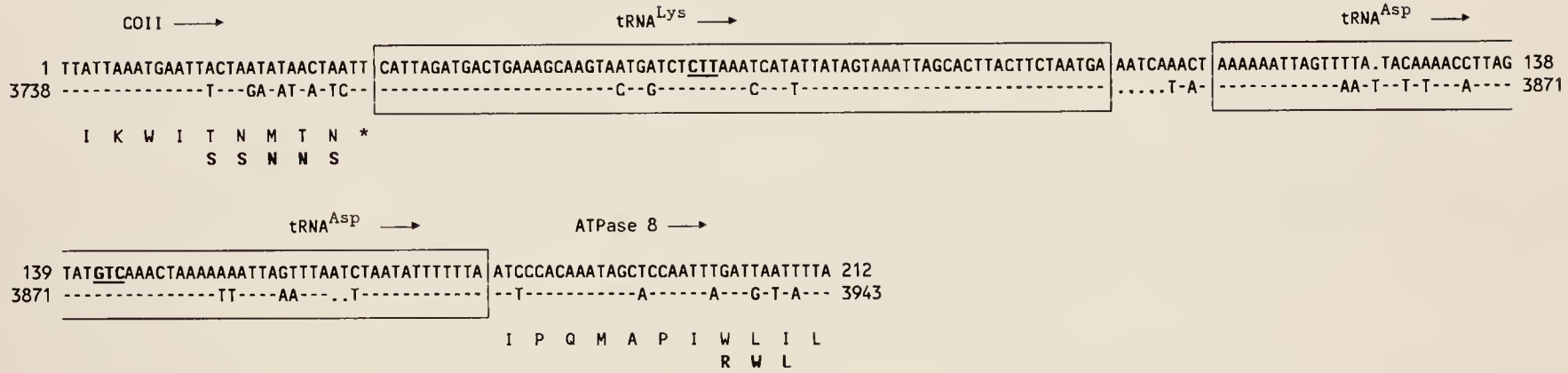

COII
tRNA Lys
tRNA Asp
1 TTATTAAATGAATTACTAATATAACTAATT CATTAGATGACTGAAAGCAAGTAATGATCTCTTAAATCATATTATAGTAAATTAGCACTTACTTCTAATGA AATCAAACT AAAAAATTAGTTTTA.TACAAAACCTTAG 138
3738 ----------T---GA-AT-A-TC-- ----------C--G---------C---T---------- .....T-A- ----------AA-T--T-T---A---- 3871
I K W I T N M T N *
S S N N S
tRNA Asp
ATPase 8
139 TATGTCAAACTAAAAAAATTAGTTTAATCTAATATTTTTTA ATCCCACAAATAGCTCCAATTTGATTAATTTTA 212
3871 ----------TT----AA---..T---------- --T----------A------A---G-T-A--- 3943
I P Q M A P I W L I L
R W L

Figure 1-4f. p2R1 (Right end, Eco R1).

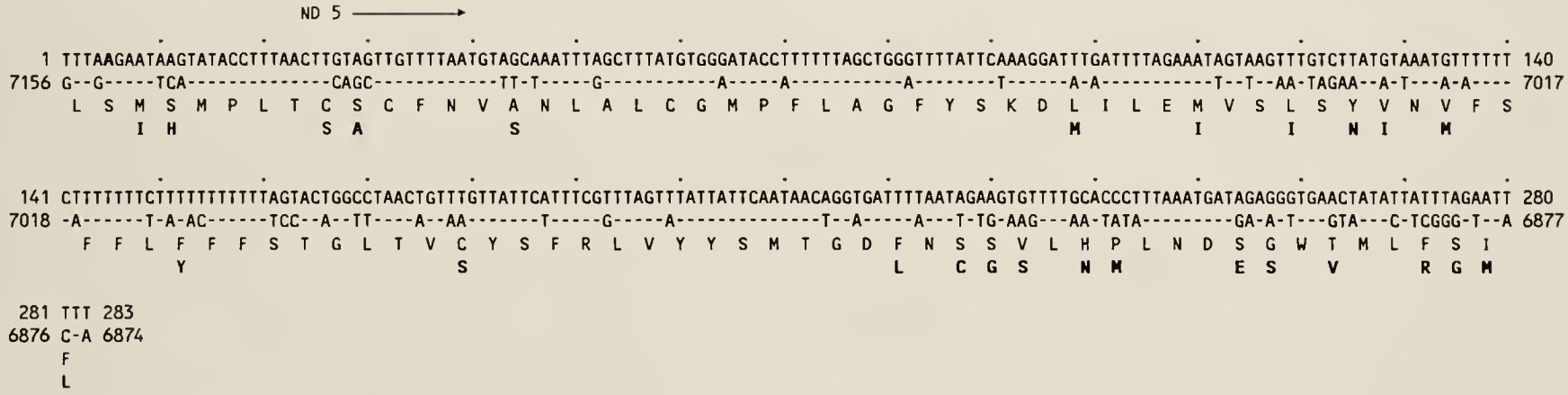

ND 5
1 TTTAAGAATAAGTATACCTTTAACTTGTAGTTGTTTTAATGTAGCAAATTTAGCTTTATGTGGGATACCTTTTTTAGCTGGGTTTTATTCAAAGGATTTGATTTTAGAAATAGTAAGTTTGTCTTATGTAAATGTTTTTT 140
7156 G--G-----TCA-------------CAGC-----------TT-T-----G----------A-----A----------A--------T------A-A----------T--T--AA-TAGAA--A-T---A-A---- 7017
L S M S M P L T C S C F N V A N L A L C G M P F L A G F Y S K D L I L E M V S L S Y V N V F S
I H S A S M I I N I M
141 CTTTTTTTCTTTTTTTTTTTAGTACTGGCCTAACTGTTTGTTATTCATTTCGTTTAGTTTATTATTCAATAACAGGTGATTTTAATAGAAGTGTTTTGCACCCTTTAAATGATAGAGGGTGAACTATATTATTTAGAATT 280
7018 -A------T-A-AC------TCC--A--TT----A--AA-------T-----G-----A-------------T--A-----A---T-TG-AAG---AA-TATA---------GA-A-T---GTA---C-TCGGG-T--A 6877
F F L F F F S T G L T V C Y S F R L V Y Y S M T G D F N S S V L H P L N D S G W T M L F S I
Y S L C G S N M E S V R G M
281 TTT 283
6876 C-A 6874
F
L

Figure 1-4g. p2R1 (left end, Eco R1).

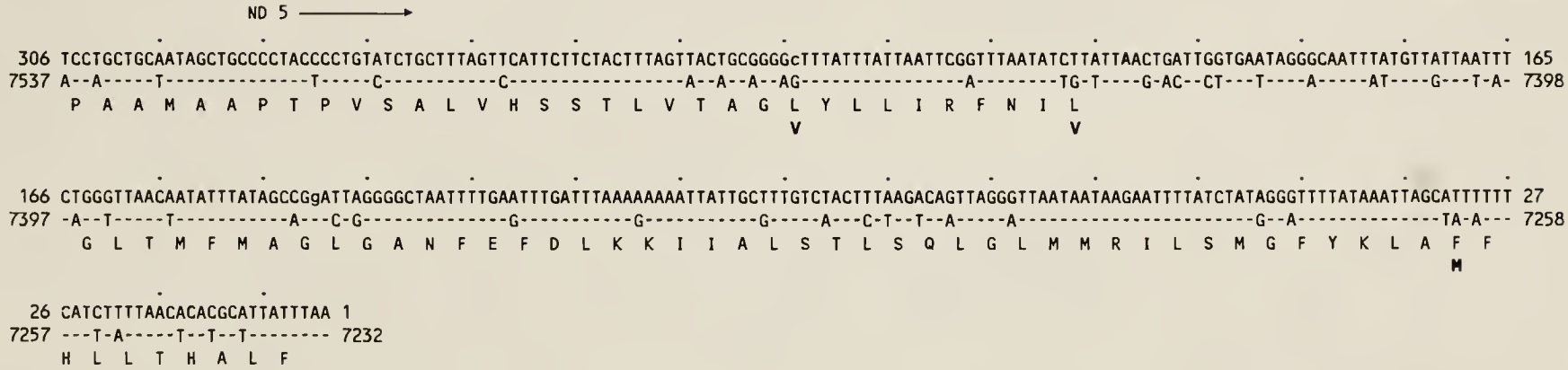

ND 5
306 TCCTGCTGCAATAGCTGCCCCTACCCCTGTATCTGCTTTAGTTCATTCTTCTACTTTAGTTACTGCGGGGcTTTATTTATTAATTCGGTTTAATATCTTATTAACTGATTGGTGAATAGGGCAATTTATGTTATTAATTT 165
7537 A--A-----T-------------T-----C----------C----------------A--A--A--AG---------------A--------TG-T----G-AC--CT---T----A-----AT----G---T-A- 7398
P A A M A A P T P V S A L V H S S T L V T A G L Y L L I R F N I L
V V
166 CTGGGTTAACAATATTTATAGCCGgATTAGGGGCTAATTTTGAATTTGATTTAAAAAAAATTATTGCTTTGTCTACTTTAAGACAGTTAGGGTTAATAATAAGAATTTTATCTATAGGGTTTTATAAATTAGCATTTTTT 27
7397 -A--T-----T----------A---C-G----------G----------G----------G-----A---C-T--T--A-----A----------G--A----------TA-A--- 7258
G L T M F M A G L G A N F E F D L K K I I A L S T L S Q L G L M M R I L S M G F Y K L A F F
M
26 CATCTTTTAACACACGCATTATTTAA 1
7257 ---T-A-----T--T--T-------- 7232
H L L T H A L F

Figure 1-4h. p1RN2 (right end, Eco R1).

ND 4
244 AAGATTGTGGTGATTTTTATTATGTTCAGGCAATATAGCAGCTCCTCCCACTTTAAATTTATTAGGGGAAATTTCTTTATTAAATAGAATTGTGAGATGATCTTGAGTTACAATAATTATATTGGCTTTTTTATCTTTTT 105
8531 --CT--A--A----------A-A----CT----------A--G----------A----------TTCT--G--A---A--T----------AT----------C- 8392
S L W W F L L C S G N M A A P P T L N L L G E I S L L N S I V S W S W V T M I M L A F L S F F
T S A I S S
104 TTAGAGCTGCTTATTCTTTATATTTATTTGCTTATAGACAACATGGNAAAATTTTTTCTGGGGTTTATTTTTTTTCTGTAGGTACAACTCGAGAATTC 7
8391 -------A--C---A----------A-T---T---T--------T---T-A----------A----CA---AGAAG-----A--T---G----AT 8294
S A A Y S L Y L F A Y S Q H G K I F S G V Y F F S V G T T R E F
T Y S F L S S K I Y

Figure 1-4i. pAYCm1 (right end, Eco R1*).

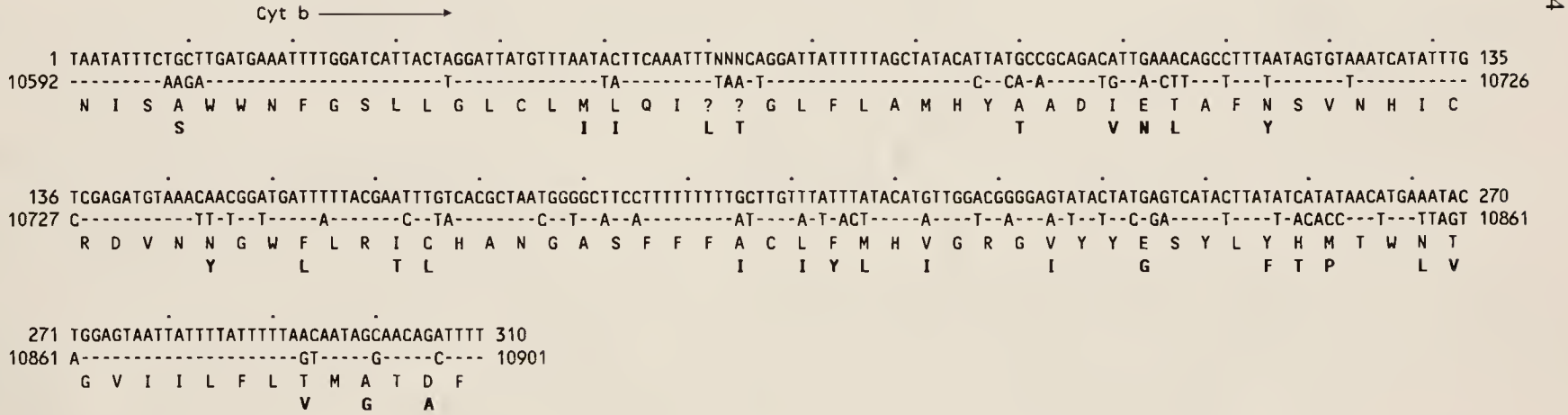

Cyt b
1 TAATATTTCTGCTTGATGAAATTTTGGATCATTACTAGGATTATGTTTAATACTTCAAATTTNNNCAGGATTATTTTTAGCTATACATTATGCCGCAGACATTGAAACAGCCTTTAATAGTGTAAATCATATTTG 135
10592 ---------AAGA----------T----------TA---------TAA-T----------C--CA-A-----TG--A-CTT---T---T-------T---------- 10726
N I S A W W N F G S L L G L C L M L Q I ? ? G L F L A M H Y A A D I E T A F N S V N H I C
S I I L T T V N L Y
136 TCGAGATGTAAACAACGGATGATTTTTACGAATTTGTCACGCTAATGGGGCTTCCTTTTTTTTTGCTTGTTTATTTATACATGTTGGACGGGGAGTATACTATGAGTCATACTTATATCATATAACATGAAATAC 270
10727 C----------TT-T--T-----A-------C--TA--------C--T--A--A---------AT----A-T-ACT-----A----T--A---A-T--T--C-GA-----T----T-ACACC---T---TTAGT 10861
R D V N N G W F L R I C H A N G A S F F F A C L F M H V G R G V Y Y E S Y L Y H M T W N T
Y L T L I I Y L I I G F T P L V
271 TGGAGTAATTATTTTATTTTTAACAATAGCAACAGATTTT 310
10861 A----------GT-----G-----C---- 10901
G V I I L F L T M A T D F
V G A

Figure 1-4j. p1RN1 (right end, Nsi I)

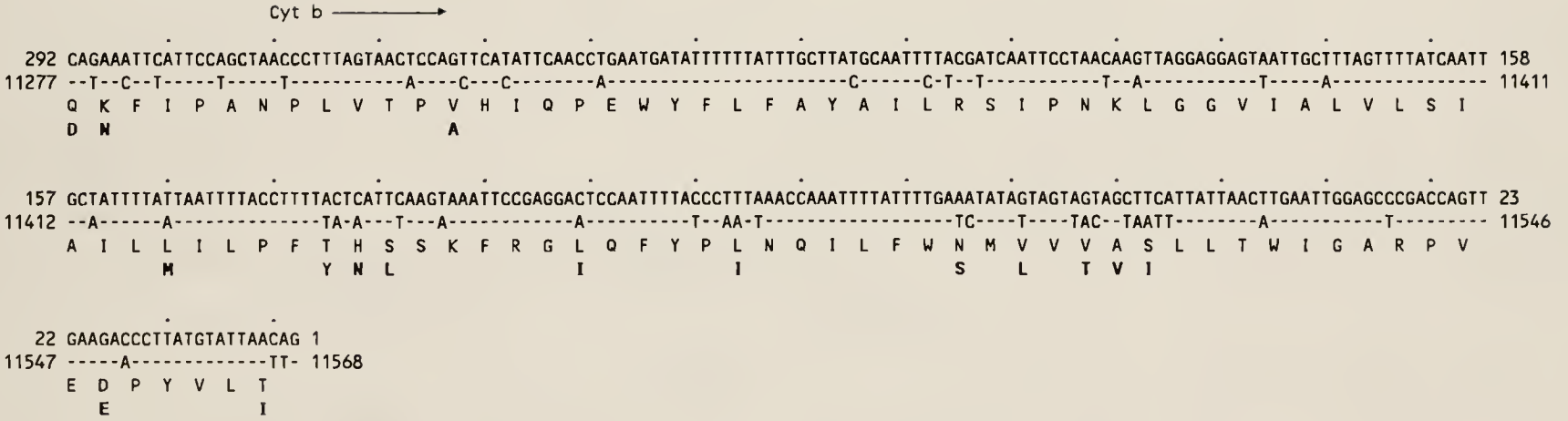

Cyt b
292 CAGAAATTCATTCCAGCTAACCCTTTAGTAACTCCAGTTCATATTCAACCTGAATGATATTTTTTATTTGCTTATGCAATTTTACGATCAATTCCTAACAAGTTAGGAGGAGTAATTGCTTTAGTTTTATCAATT 158
11277 --T--C--T-----T-----T-----T----------A----C---C--------A----------C------C-T--T----------T--A----------T-----A---------- 11411
Q K F I P A N P L V T P V H I Q P E W Y F L F A Y A I L R S I P N K L G G V I A L V L S I
D N A
157 GCTATTTTATTAATTTTACCTTTTACTCATTCAAGTAAATTCCGAGGACTCCAATTTTACCCTTTAAACCAAATTTTATTTTGAAATATAGTAGTAGTAGCTTCATTATTAACTTGAATTGGAGCCCGACCAGTT 23
11412 --A------A-------------TA-A---T---A-----------A----------T--AA-T----------------TC----T----TAC--TAATT--------A----------T--------- 11546
A I L L I L P F T H S S K F R G L Q F Y P L N Q I L F W N M V V V A S L L T W I G A R P V
M Y N L I I S L T V I
22 GAAGACCCTTATGTATTAACAG 1
11547 -----A----------TT- 11568
E D P Y V L T
E I

Figure 1-4k. p1RH1 (left end, Hind III).

ND 1
295 TTCAAGAATTTTATTTATAAGTATATTATTTGTTGTTATTTTTTTAGGGAGTGATATTTATAGACTTATATTTTTTTTTAAATTAACTTTTATTTCTTTTGTATTTATTTGAGTTCGTGGGACTTTACCTCGATT 160
11985 -G-T----------------A-----G---TG---A----------TT-----G-A-T--ATT-AT------A-G-A----------------A----------C---A--T--A--------G-- 11861
S S I L F M S M L F V V I F L G S D I Y R L M F F F K L T F I S F V F I W V R G T L P R F
A C C V F N L Y V A
159 TCGATATGATAAATTAATGTATTTAGCTTGAAAAAGATTTTTACCGTTATCTTTAAATTATTTATTTTTTTTTGTTGGGTTAAAAATTTTTTTTATTTCTTTATTATTTTAATGAATAAATTTTAGTATTA 30
11860 ---T-------------A--------------T-T------T-A--T----------------A------A-------T--------A--AT-------T-----G-...-----TTT--------.A- 11723
R Y D K L M Y L A W K S F L P L S L N Y L F F F V G L K I F F I S L L F *
C S F L I F L L F F L W I F F S K
tRNA Ser
29 AAATTAATAGAAGATTTTACTTCTTTCTT 1
11723 --G---------A---AA-.----A---- 11696
K L M E N *

Figure 1-4l. p1RH1 (right end, Hind III).

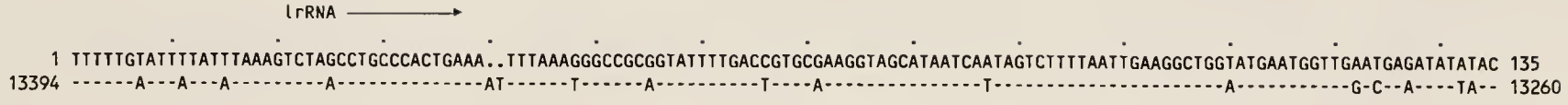


Figure 1-4m. p1RN1 (left end, Eco R1).

Figure 1-4n. p3RX2 (entire clone).

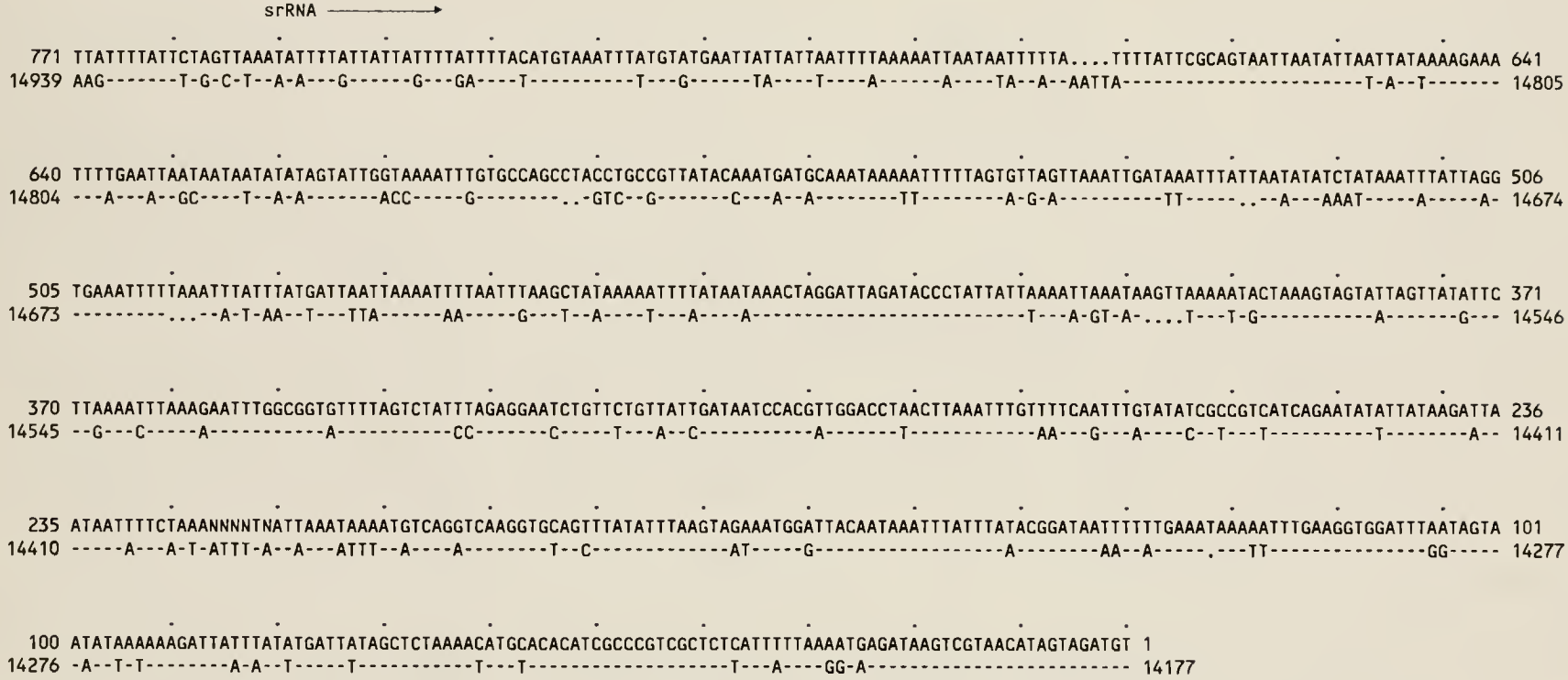


Figure 1-4o. p3X2 (entire clone).

The stop codon (TAA) is generated by an insertion (AAT) in the mosquito sequence at positions 50-48 (Figure 2-4l). Most *Drosophila* mitochondrial protein genes end in a "T" or a "TA" and the complete stop codon is generated by polyadenylation (Clary and Wolstenholm, 1985). Sequence for the end of the *CO II* gene (Figure 2-4f) was obtained for *A. quadrimaculatus.* In this case, the stop codon is not completely encoded in the mtDNA and is apparently created by polyadenylation. Therefore, it would seem likely that TAA is recognized as a stop codon in the mosquito and the ND 1 gene is, in fact, truncated.

An ATG, ATT or ATA codon occurs at the beginning of all of the *D. yakuba* mitochondrial protein genes except the *CO I* gene where the tetranucleotide ATAA appears to serve as the translation initiation codon (Clary and Wolstenholme, 1983). Limited sequence was obtained from the 5' end of the *ATPase 8* gene of *A. quadrimaculatus* (Figure 2-4f). In this case, either an ATC triplet (positions 180-182) functions as the initiation codon or translation begins seven bases downstream at ATA. In *Drosophila melanogaster*, the ATC codon appears internally and is not used to initiate translation (Garesse, 1988).

The *tRNA* genes are extremely well conserved between *A. quadrimaculatus* and *Drosophila* (Table 2-1; Figures 2-4a, 2-4f and 2-4n). In general, the mosquito tRNAs are slightly longer than those in *Drosophila* and have fewer bases separating the genes (Figure 2-4a).

The A+T-rich region has little or no detectable homology to the *Drosophila* sequence, aside from that due to their highly skewed base composition (Table 2-1, Figure 2-5). When the *D. yakuba* A+T rich region was randomly shuffled and tested

Figure 2-5. Sequence of *Anopheles quadrimaculatus* mtDNA A+T-rich region.

(a) Sequence of p3RU1 (right end, *Xba* I); (b) Sequence of p3X1 (right end, *Xba* I).

Note: These sequences do not show significant similarity to *D. yakuba*.

```
    TTAATATCACCTATTTAATATAAATTATTAATATATATAAATAAATATATTAAATATAAATTATTTATATATATATTAATAATTTATAATAAATAAATAAAATCTCCTTATTAATTTAATGTATAGGTCCATTTTTTAAA
  1 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 140
    AATTATAGTGGATAAATTATATTTAATAATTATATATATTTATTTATATAATTTATATTTAATAAATATATATATAATTATTAAATATTATTTATTTATTTTAGAGGAATAATTAAATTACATATCCAGGTAAAAAATTT

    AATGTAATAATAAATTAAAAATTTTTAAANNATTAANNAAAATCAAAAT
141 ---------+---------+---------+---------+--------- 189
    TTACATTATTATTTAATTTTTAAAAATTTNNTAATTNNTTTTAGTTTTA
```

Figure 1-5a. p3RU1 (right end, Stu I).

```
    CCTcATTTAATTTGTTNNCCCTAATTTTTTTTTTTTTTTTTTAATATAATTTATTTAAATTAAATTATATTTTAAATAAAACTTTATATATTTTATATATATATAAATATTACAATAAATTAATTTATATATTAATAATAT
  1 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 140
    GGAgTAAATTAAACAANNGGGATTAAAAAAAAAAAAAAAAAAATTATATTAAATAAATTTAATTTAATATAAAATTTATTTTGAAATATATAAATATATATATATTTATAATGTTATTTAATTAAATATATAATTATTATA

    AATTTATATATTAATAATATTTAAATAATTAATTTAAATTATAATTATATATATATAATATATTTATATAAA
141 ---------+---------+---------+---------+---------+---------+---------+-- 192
    TTAAATATATAATTATTATAAATTTATTAATTAAATTTAATATTAATATATATATATTATATAAATATATTT
```

Figure 1-5b. p3RX1 (right end, Xba I).

for homology to the *A. quadrimaculatus* A+T rich region, the homology was as good as that of the unshuffled sequence. The similarity statistics calculated for the best alignment of the unshuffled *Drosophila* sequence and clones p3RU1 and p3X1 were 71.3% and 69.7%, respectively. The mean similarities for alignments of 20 randomized *Drosophila* sequences were 72.8 ± 2.3% for p3RU1 and 73.0 ± 3.7% for p3X1.

The mtDNAs of *D. yakuba* and *A. quadrimaculatus* differ in size by about 0.7 kb (*D. yakuba* mtDNA is 16 kb and *A. quadrimaculatus* is 15.3 kb). Since all genes were mapped to the same locations in both species, the length difference of the mitochondrial genomes is attributed to variation in length of the A+T rich origin of replication.

Discussion

Animal mtDNAs are small molecules and should be easy to clone in bacteria. However, many groups have found it difficult or impossible to clone an entire mitochondrial genome in one piece (W. Hauswirth, personal communication). Many DNAs that have been found to be "unclonable" contain segments capable of forming 2° and 3° structures such as cruciforms (due to the presence of inverted repeats) (Erlich, 1989) or Z-DNA (within stretches of alternating purine-pyrimidine residues) (Santella et al., 1981). Sequencing has shown that the mtDNA of *A. quadrimaculatus,* like *Drosophila,* is extremely A+T-rich (about 20% G+C). Long adjacent stretches of A's and T's and AT repeats are frequently encountered (particulary in the tRNA and rRNA genes and in the region of the replication origin). It has been demonstrated that cloning the mitochondrial genome of mosquitoes in bacteriophage can be achieved by

cleaving the genome into several pieces. This success may be due to: (1) the interruption of segments that have a propensity to form non-standard structure; (2) the fortuitous destruction, by the cloning strategy, of a single inserted gene that may be lethal to the bacterial host; (3) differences in host/vector systems, or; (4) the separation of several genes that interact to make the mtDNA unstable or lethal to bacteria. The three *Eco* RI sites used in the cloning interrupt the *lrRNA* gene, the *CO II* gene, and the *ND 5* gene (Figure 2-2).

In order to subclone the bacteriophage inserts into plasmids, it was necessary to cut the mtDNA into still smaller fragments. Apparently there are even greater restraints on DNA inserted into plasmids than on DNA inserted into bacteriophage. This is probably because bacteriophage are less dependent on the proper working of the cell, since after replicating they lyse the cell and infect another. Plasmids, which are not capable of infection, will be sensitive to any sequence that disrupts the long term functioning of the cell.

Identification of the plasmid subclones containing part of the *lrRNA* gene was difficult because they restored the lac^+ phenotype to recombinant plasmids. The multiple cloning site of plasmid vector pUC 19 is located within the *E. coli* β-galactosidase α-peptide gene. The multiple cloning site in pUC 19 interrupts the β-galactosidase gene 18 base pairs downstream from the translation start site of the α-peptide. Therefore, insertion of DNA at the multiple cloning site usually results in interruption of the α-peptide, producing lac^- colonies, which can be identified on media containing the appropriate chromogenic substrate. All subclones having the aberrant

phenotype contained the 0.6 kb *Eco* RI/ *Xba* I fragment which consisted of a *tRNA* and the 5' end of the *lrRNA* gene (see Figure 2-4n).

Restoration of gene activity occasionally occurs due to the insertion of a fragment containing a completely open reading frame but sequencing has shown that such an open reading frame does not exist in subclone p3RX2. Alternatively, restoration of β-galactosidase activity would be possible if sequences exist within this stretch of mtDNA that act as a strong promoter in bacteria, contain a transcriptional start site, and contain an in-frame translation start site.

Bacterial promoters possess two highly conserved sequences. A sextamer (TTGACA) occurs about 35 bases upstream of the transcription start site. This sequence is recognized but is not bound by RNA polymerase. The Pribnow box (TATAAT) is located 5-9 bases upstream from the transcription startpoint and because of its high A+T content is probably involved with initial melting of DNA into single strands. The transcription start site is not highly conserved (Lewin, 1985).

There are several sequences in p3RX2 that are similar to the conserved -35 sextamer of bacterial promoters. Two of these sequences occur within three bases of each other (TTGATA at positions 551-556 and TTGTTA at positions 560-565) (Figure 2-4n). One of the reading frames opens at position 633. The stretch of DNA (about 70 bases) between the "-35-like" sequences and this open reading frame is extremely A+T-rich. The distance between the conserved sextamer and the transcription initiation site, however, is about half this length in bacterial promoters (35 bases). A translation start codon (ATG) is present in this reading frame at positions 650-652. The DNA sequence obtained for subclone p3RX2, unfortunately, did not include the *Eco* RI

cloning site. Therefore, it is not known if the insert sequence is in frame with the remainder of the α-peptide sequence contained in the plasmid vector and, as a result, restoration of β-galactosidase function in this clone remains a mystery.

Even though the A+T-rich region has been successfully cloned in *Drosophila yakuba,* several groups have failed to clone the A+T-rich origin of replication and flanking sequences of *D. melanogaster* mtDNA either in bacteriophage or in plasmids (Mason and Bishop, 1980; Garesse, 1988). The 5.6 kb Eco RI fragment containing this region of *A. quadrimaculatus* mtDNA was readily cloned in bacteriophage. However, when initial attempts were made to subclone this fragment into plasmids, a rearranged subclone was obtained (pAYCm3). This subclone had a 0.7 kb insertion located within the 2.7 kb *Hind* III fragment which contains the A+T-rich region. Subsequently, several additional subclones of this fragment were isolated and all had deletions within this same region. Stable, unrearranged plasmid subclones (p3X1 and p3RU2) were obtained by subcloning DNA containing the A+T-rich region as small fragments. Unrearranged subclones that contained the 4.2 kb *Eco* RI/ *Stu* I fragment from AqAm3, however, were not recovered. Apparently only plasmids having about 1.5 kb or less DNA flanking the A+T-rich region were stable in bacteria.

The genetic organization of the mtDNA of *A. quadrimaculatus* is similar to that of *Drosophila.* All of the protein coding sequences that have been localized were in the same positions in both species. Considerable homology is seen between protein coding regions, though no homology was seen between the A+T-rich origins of replication. Six tRNAs were identified ($tRNA^{Lys}$ and $tRNA^{Asp}$ from subclone p2R1, $tRNA^{Ile}$, $tRNA^{Gln}$ and $tRNA^{fMet}$ from p3RU1 and $tRNA^{Val}$ from clone p3RX2). These *tRNA* genes were located

in the same positions as in *Drosophila.* In *Aedes albopictus* two of the six contiguous *tRNA* genes located between *ND 5* and *ND 3* are inverted with respect to the *Drosophila* map (Dubin et al., 1986). It is not known if this reshuffling of *tRNAs* is also present in *A. quadrimaculatus,* since sequence information has not yet been obtained from this region.

The cloning of the complete mtDNA genome from a mosquito should greatly improve the accuracy of mtDNA restriction analysis. This powerful tool is now being widely used to study the population genetics of wild mosquitoes, and with the use of mosquito mtDNA as a probe, can be used to study the mtDNA of individual mosquitoes (Cockburn and Seawright, 1988). The *A. quadrimaculatus* species complex consists of at least four sibling species: species A, species B (Lanzaro et al., 1988; Kaiser et al., 1988a), species C (Kaiser et al., 1988b; Narang et al., 1989a), and species D (Narang et al. 1989b). The restriction profiles of the mtDNA of the four species of this complex have been determined, and this information will be used to estimate the degree of diversity between different populations and to find diagnostic restriction sites for identifying individual insects.

CHAPTER III

MITOCHONDRIAL AND RIBOSOMAL DNA VARIATION AMONG MEMBERS OF THE *Anopheles quadrimaculatus* SPECIES COMPLEX

Introduction

The use of restriction endonucleases to measure mitochondrial DNA (mtDNA) sequence relatedness in natural populations has, over the past few years, become an important tool for population biologists. Genetic variation in natural populations has traditionally been studied by electrophoretic analysis of the protein products of nuclear genes. Identification of enzyme variants (isozymes) by electrophoresis has proven to be a valuable and versatile technique but this method is limited by certain fundamental characteristics of the nuclear genome. Since the nuclear genome is highly recombinant and gene flow between populations (via migration or dispersal) can lead to genetic homogenization, the reconstruction of historical relationships within a species becomes difficult and subtle population structuring can be missed (Avise et al, 1979; Saunders et al., 1986). Animal mitochondrial DNA, on the other hand, is maternally inherited, haploid, and nonrecombinant (Dawid, 1972; Hutchinson et al., 1974; Lansman et al., 1983; Gyllensten et al., 1985). Therefore, the mitochondrial genotype (as determined by restriction site maps) can be used as a marker to trace maternal lineages and the genetic relationships among individuals can be established. Also, mtDNA analysis may be a more sensitive way to detect population structure than the traditional electrophoretic

analysis of nuclear enzyme-coding loci since mtDNA is less susceptible to interpopulation genetic homogenization due to gene flow and more susceptible to population bottlenecks because of its maternal mode of transmission and haploid copy number (Birky et al. 1983; Wilson et al., 1985).

Until recently, the mosquito, *Anopheles quadrimaculatus* (Say), which is widely distributed over the eastern United States, was considered to be a single species. Results of hybridization, cytogenetic and electrophoretic studies, however, have lead to the discovery of at least four sympatric sibling forms, species A, B (Lanzaro 1986; Lanzaro et al., 1988; Kaiser et al., 1988a), C (Kaiser et al., 1988b, Narang et al., 1989a), and D (Narang et al., 1989b). Over the past three years, extensive electrophoretic data have been collected for 147 natural populations of *A. quadrimaculatus.* The populations studied were located primarily in the southeastern United States. Diagnostic allozyme loci were described for each sibling species and electrophoretic keys have been published (Narang et al., 1989a,b).

This study reports the results of the analysis of mtDNA restriction site polymorphisms in seven mixed populations (i.e. more than one sibling form present) of *A. quadrimaculatus.* The primary objective of this study was to investigate the feasibility of using mtDNA and ribosomal DNA (rDNA) restriction patterns for the identification of individuals and species of mosquitoes. Since the rDNA profiles in conjunction with isozyme analysis provided markers for the nuclear genomes of individual mosquitoes, interspecific hybridizations and introgression of mtDNA between species could potentially be detected. In addition, the mtDNA restriction site data were used to evaluate the genetic relationships among the four sibling species of the *A. quadrimaculatus* complex.

Materials and Methods

Collection sites

Adult *A. quadrimaculatus* from seven natural populations (four from Florida, one from Alabama and two from Mississippi) (Table 3-1), were collected by vacuum aspiration from tree holes or artificial resting stations. The taxonomic key of Darsie and Ward (1981) was used to identify the adult mosquitoes. All confirmed specimens were frozen in liquid nitrogen for transport back to the laboratory where they were stored at -80° C until analyzed.

Isozyme analysis

Frozen mosquitoes were placed on ice and decapitated. Horizontal starch gel electrophoresis was performed on protein homogenates obtained from the head of each mosquito. Samples from 50 individuals were analyzed for each of the seven populations surveyed. Electrophoresis was done according to Steiner and Joslyn (1979) with minor modifications as described by Narang et al. (1989a). Gels were stained for enzymes that were known to be diagnostic for the four sibling species of the *A. quadrimaculatus* complex (Lanzaro, 1986; Narang et al., 1989a,b). A single locus or combinations of loci have been considered to be diagnostic if the probability of correct assignment of a mosquito to a species was >99% (Ayala and Powell, 1972). For most populations, two loci each of *isocitrate dehydrogenase* (*Idh-1* and *Idh-2*), *hydroxy acid dehydrogenase* (*Had-1* and *Had-3*) and *glutamate oxaloacetate transaminase* (*Got-1* and *Got-2*) were scored as well as one locus for *phosphoglucose isomerase* (*Pgi-1*) and *malic enzyme* (*Me-1*) (Table 3-1).

Table 3-1. Species composition and enzyme loci analyzed for seven natural populations of *Anopheles quadrimaculatus.*

POPULATION	SPECIES[a] (%) A	B	C	D	ENZYME LOCI ANALYZED
Lake Octahatchee Hamilton Co., FL (LOC)	9.3	90.7	0	0	Idh-1, Idh-2, Had-3, Pgi-1, Me-1
Dead Lake Calhoun Co., FL (DLK)	62.2	37.8	0	0	Idh-1, Idh-2, Had-3, Pgi-1, Me-1, Got-1, Got-2
Bear Bay Swamp Dixie Co., FL (BBS)	0.8	0	99.2	0	Idh-1, Idh-2, Pgi-1, Me-1, Got-1, Got-2
Lolly Creek Dixie Co., FL (LCK)	32.3	17.7	50.0	0	Idh-1, Idh-2, Had-3, Pgi-1, Me-1
Wheeler Reservoir Limestone Co., AL (WHR)	45.2	54.8	0	0	Idh-1, Idh-2[b], Had-3, Pgi-1, Me-1, Got-1, Got-2
Pickwick Reservoir Tishomingo Co., MS (PIC)	21.2	5.1	0	73.7	Idh-1, Idh-2, Had-1, Had-3 Pgi-1, Me-1
Noxubee Wild Life Refuge Noxubee Co., MS (NOX)	31.6	68.4	0	0	Idh-1, Idh-2[b], Had-3, Pgi-1, Me-1

[a] Species composition was based on unpublished electrophoretic data.

[b] The majority of individuals were analyzed for Idh-1 and Idh-2 only.

DNA Analysis

Total genomic DNA was extracted from the remainder of each decapitated mosquito by the procedure of Cockburn and Seawright (1988). The individual DNA preparations were stored as ethanol precipitates at -20° C. Nucleic acid precipitates were recovered by centrifugation for 5 min. in a microfuge. The pellets were washed twice with 70% ethanol, air dried and redissolved in 25 µl TE (10 mM tris HCl, 1 mM EDTA, pH 8.0) containing preboiled RNase A at a concentration of 1 µg/ml. After incubation at room temperature for 30 min, restriction enzyme digestions were performed according to protocols provided by the manufacturer (Bethesda Research Laboratories). The total genomic DNA preparations from individual mosquitoes were divided into five portions. Each of four aliquots was cleaved separately with a different restriction enzyme (*Ava* I, *Hind* III, *Pst* I or *Pvu* II). Restriction enzyme digestions were done at 37° C for 2 hr with a ten-fold excess of enzyme. The remaining one-fifth of the mosquito DNA preparation was retained in case of failure of one of the cleavage reactions or for further analysis. The digested DNA samples were electrophoresed in 0.8% agarose gels and Southern blots were prepared according to established procedure (Berger and Kimmel, 1987).

Southern blots were first hybridized to a mixture of ^{32}P-radiolabeled mtDNA clones (p1RN1, p1RN2, p2R1, p3RU1 and p3RU2). These plasmid clones, described in Chapter II, comprise the entire mitochondrial genome of *A. quadrimaculatus* species A. Nick translations were done with purchased reagent kits and according to the manufacturer's instructions (Bethesda Research Laboratories). Nitrocellulose filters were

blocked and hybridizations were performed as in Cockburn and Seawright (1988). Autoradiography was a standard technique (Berger and Kimmel, 1987).

Nitrocellulose filters were stripped of bound mtDNA probe by brief immersion in boiling water followed by several washes in warm water. The filters were blocked, as before, and rehybridized to an rDNA clone, QB1, from *A. quadrimaculatus,* species B. The QB1 clone was isolated from a total genomic DNA library in bacteriophage λ vector EMBL3A (Cockburn and Mitchell, 1989). This mosquito clone was homologous to pBC2, a pBR322 clone that contained the entire 8.4 kb rDNA repeat unit from the fungus gnat, *Sciara coprophila* (Renkawitz and Gerbi, 1979). Since the QB1 clone has not been extensively characterized, whether or not it includes the entire rDNA cistron of *A. quadrimaculatus* species B is unknown.

Restriction Site Mapping of the Mitochondrial Genomes

Restriction sites for *Ava* I, *Hind* III, *Pst* I and *Pvu* II were mapped for the mtDNA of one isofemale line of each of the four sibling species. The cleavage-site maps were derived for the most common restriction patterns of each species and the positions of rare sites were deduced from these maps.

Total genomic DNA was isolated as described above and complete single and double digests were done for all combinations of the four restriction enzymes listed above. The cleavage reactions were electrophoresed in 0.8 % agarose gels and Southern blots were made and hybridized to the mtDNA plasmid clones as previously described.

Cleavage-site mapping was based on the analysis of fragment sizes estimated from the electrophoretic patterns obtained for the single and double digests. The *Hind* III fragments of phage λ DNA were used as size standards and a Lotus program (FRAGMENT) was used to calculate fragment sizes based on the distance migrated. The FRAGMENT program was obtained from M. Q. Benedict (USDA-ARS, Insects Affecting Man and Animals Research Laboratory).

Estimation of Genetic Distance

Sequence divergence, expressed as the frequency of nucleotide substitutions (δ), was estimated for the four sibling species of *A. quadrimaculatus* from the number of sites mapped and the number of sites shared (S) in the genomes according to formulae (8) and (10) of Nei and Li (1979). Sample variances of S and δ were estimated by formulae (19) and (20) of Nei and Tajima (1981). Pairwise comparisons were made for all haplotypes. Interspecific distance calculations for each enzyme class were combined after weighting by haplotypic frequency and the proportion of the total number of observed restriction sites contributed by each group of enzymes across all species. Genetic distances were used to construct an unrooted phylogenetic tree based on the neighbor-joining method of Saitou and Nei (1987). Tree construction was accomplished with the aid of a FORTRAN program that was kindly furnished by Dr. Masatoshi Nei (University of Texas).

Results

Species-Specific Restriction Site Polymorphisms

Mitochondrial DNA analysis

Complete isozyme and mtDNA restriction data were obtained for 288 individual mosquitoes out of a total of 350 analyzed. Comparable sample sizes were obtained for all populations except Pickwick Reservoir (PIC). Suitable isozyme and restriction data were obtained from only 40% (20/50) of the individuals analyzed from Pickwick and the poor quality of these results was attributed to sample degradation.

The mitochondrial DNA restriction patterns for each enzyme were designated by Arabic numbers (e.g. *Ava*-1, *Ava*-2 etc.). These designations indicated the order in which the various polymorphic restriction patterns were noted on autoradiograms. For all species combined, 7 different patterns were observed for *Ava* I, 10 for *Hind* III, 6 for *Pst* I and 4 for *Pvu* II. The sizes of the fragments observed or inferred in these alternative restriction patterns are shown in Table 3-2 and pattern distribution by species and population is presented in Table 3-3. Restriction site maps for each sibling species are included in Figure 3-1. In these maps, variable restriction sites for each enzyme were denoted by lower case letters (eg. *Ava*-a, *Ava*-b, etc.).

All four of the restriction enzymes used in this study produced one major electrophoretic pattern for each species with minor variants occurring at frequencies that ranged from 1-20% (Table 3-3). For *Ava* I, a three-banded pattern (*Ava*-2) was most common in species A and C while a two-banded variant (*Ava*-1) was predominant in both species B and D (Figures 3-2a and 3-3a). Minor variants were shared in about 1% of species A and B (*Ava*-3), and species A and C (*Ava*-5). One species C individual

Table 3-2. Fragment sizes observed in alternative mtDNA restriction patterns of *Anopheles quadrimaculatus* sibling species.

Enzyme	Variant	Fragment sizes (kb)
Ava I	1	12.1; 3.3
	2	9.4; 3.3; 2.7
	3	9.4; 3.3; (1.35)[a]; (1.35)
	4	9.4; 3.1; 2.7; 0.3
	5	9.4; 6.0
	6	12.7; 2.7
	7	15.4
Hind III	1	6.3; 2.7; 2.0; 1.9; 1.7; 0.5; 0.3
	2	6.3; 2.7; 2.3; 1.9; 1.7; 0.5
	3	6.3; 2.7; 3.6; 1.9; 1.7; 0.5; 0.3
	4	4.4; 2.7; 2.0; 1.9; 1.7; 0.5; 0.3
	5	3.4; (2.7); (2.7); 2.3; 1.9; 1.7; 0.5
	6	3.9; 3.4; (2.7); (2.7); 1.9; 0.5
	7	5.6; (2.7); (2.7); 1.9; 1.7; 0.5
	8	6.3; 3.9; 2.7; 1.9; 0.5
	9	(2.7); (2.7); (2.7); 2.2; 1.9; 1.7; 0.9; 0.5
	10	(2.7); (2.7); (1.9); (1.9); (1.7); (1.7); 0.5
Pst I	1	7.65; 4.35; 3.3
	2	15.4
	3	(7.65); (7.65)
	4	10.95; 4.35
	5	10.3; 5.0
	6	6.95; 4.35; 3.3; 0.7
Pvu II	1	11.8; 3.5
	2	6.2; 5.6; 3.5
	3	9.7; 5.6
	4	15.4

[a]Inferred fragments are enclosed in parentheses.

Table 3-3. Number of individual mosquitoes observed for alternative restriction patterns for four enzymes in natural populations of *Anopheles quadrimaculatus* sibling species.

Enzyme and Variant	SPECIES BY POPULATION								
	LOC		DLK		WHR		NOX		
	A	B	A	B	A	B	A	B	D
Ava I									
1...		35		12		21		35	1
2...	7		29		26		9		
3...		1			1				
4...									
5...									
6...									
7...		1						1	
Hind III									
1...		36		12		21		33	
2...	6		28		27		8		
3...		1						2	
4...								1	
5...	1								1
6...									
7...									
8...			1				1		
9...									
10...									
Pst I									
1...		19		8		21		36	
2...	3	2	23		20		9		
3...	2		4		6				
4...	2	16	1	4	1				1
5...			1						
6...									
Pvu II									
1...	7	37	26	12	27	21	9	36	
2...									1
3...									
4...			3						

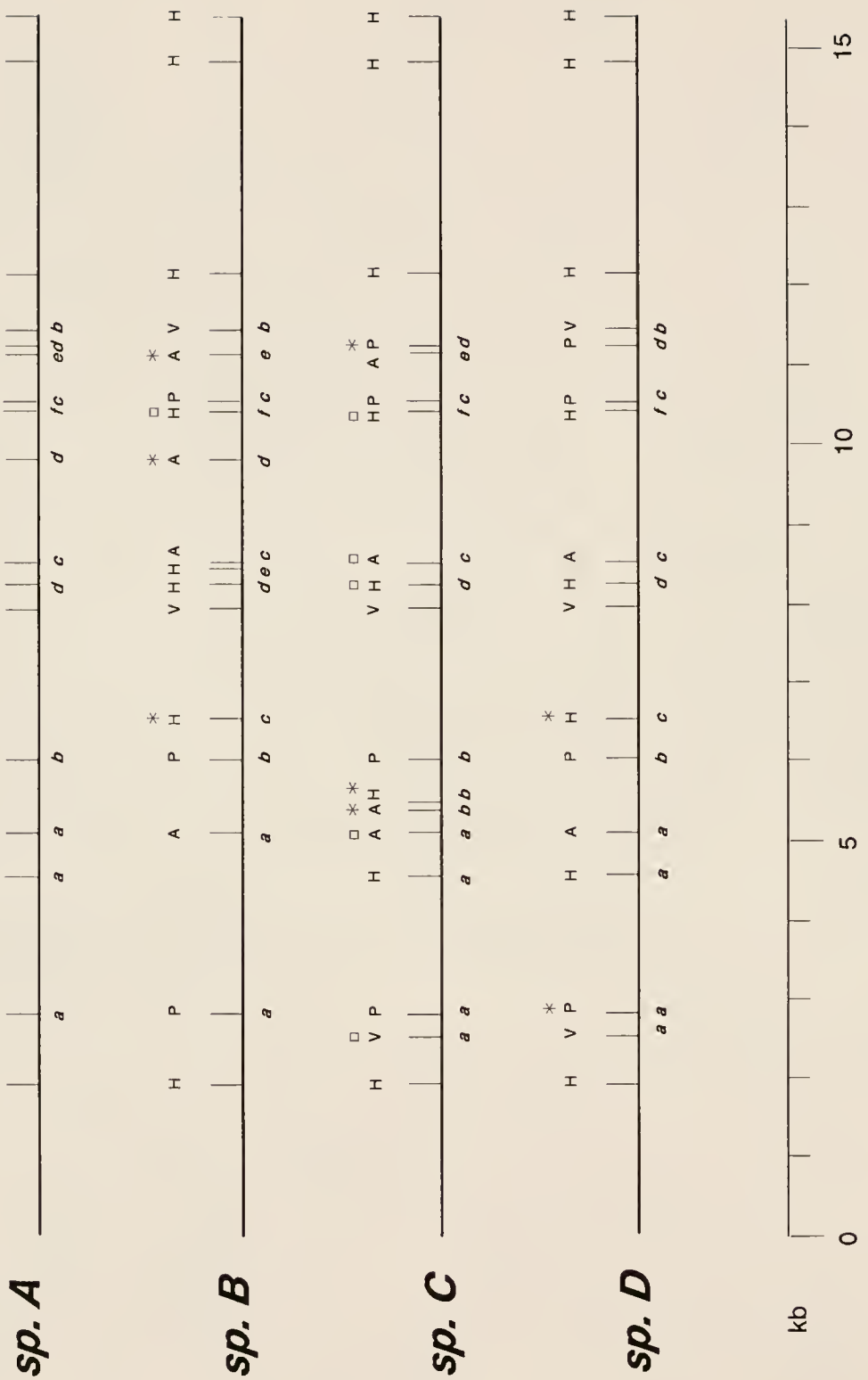
Table 3-3. ---- continued.

SPECIES BY POPULATION

	PIC			BBS		LCK			TOTAL			
	A	B	D	A	C	A	B	C	A	B	C	D
Ava I												
1...		1	20		1		2			106	1	21
2...	2			1	44	6		28	80		72	
3...									1	1		
4...								1			1	
5...					1	1			1		1	
6...					1						1	
7...										2		
Hind III												
1...		1					2			105		
2...	2			1		7			79			
3...										3		
4...										1		
5...			19		44			28	1		72	20
6...								1			1	
7...					2						2	
8...									2			
9...					1						1	
10...			1									1
Pst I												
1...		1			39			26		85	65	
2...	2		1	1	1	5			63	2	1	1
3...					5	2		2	14		7	
4...			19		2		2		4	22	2	20
5...									1			
6...								1			1	
Pvu II												
1...	2	1		1	7	2			79	109		
2...			20									21
3...					46			28			74	
4...					1			1	3		2	

Figure 3-1. Restriction maps of the mitochondrial genomes of *Anopheles quadrimaculatus* sibling species.

Note: A=*Ava* I, H=*Hind* III, P=*Pst* I, V=*Pvu* II. Open boxes and asterisks denote restriction sites that are rarely absent or rarely present, respectively. Variable sites are designated by lower case letters.


sp. A
sp. B
sp. C
sp. D
kb
0
5
10
15

from the Bear Bay Swamp population had the *Ava*-1 pattern which was prevalent in species B and D (Table 3-3).

Approximately 95% of species C and D and one species A individual from the Lake Octahatchee population had the *Hind*-5 mtDNA pattern (Figures 3-2b and 3-3b). The majority (97.6%) of species A mosquitoes, however, possessed a pattern that was unique to that species (*Hind*-2). All of the species B variants observed were unique to that species because of the presence of the diagnostic *Hind*-e site which cleaves the 2.3 kb *Hind* III fragment common to the other 3 species (Figure 3-1).

Slight variation in the mobility of the *Hind* III fragment containing the A+T-rich replication origin was noted in a few of the species A and species B mosquitoes analyzed. The size of this DNA fragment was usually 2.7 kb in all four sibling species. The species B individual in Figure 3-2b (third lane), however, had a 2.9 kb *Hind* III fragment. The length variation observed for this fragment was typically on the order of a few hundred base pairs. Because of the weak hybridization to the A+T-rich region and the small differences in mobility between these size variants, all individuals were not classified with regard to this length polymorphism. This variation was also not included when the maps in Figure 3-1 were developed.

It should be noted that the 0.5 kb *Hind* III fragment common to all of the *A. quadrimaculatus* siblings and the 0.3 kb *Hind* III fragment present in species B were observed on autoradiograms only if larger quantities of total DNA were loaded (at least 0.5 μg, which is about one-half of one individual mosquito DNA preparation) and exposure times were lengthened (Figure 3-2c).

Figure 3-2. Variable mitochondrial DNA restriction patterns for *Anopheles quadrimaculatus* sibling species.

(a) *Ava* I mtDNA restriction patterns; (b) *Hind* III mtDNA restriction patterns; (c) Overexposure of autoradiogram shown in panel (b). Upper arrow marks the position of the 0.5 kb *Hind* III fragment common to all four sibling species. Lower arrow denotes the 0.3 kb *Hind* III fragment present in species B; (d) *Pst* I mtDNA restriction patterns; (e) *Pvu* II mtDNA restriction patterns; (f) *Cla* I mtDNA restriction patterns.

Note: S=*Hind* III fragments of bacteriophage λ. Lanes containing sibling species were loaded in the same order from left to right for all panels. A=species A from the ORLANDO laboratory strain; A=species A from Lake Panasofkee, Florida; B=species B from Wheeler Reservoir; C=species C from Bear Bay Swamp; D=species D from Pickwick Reservoir. Pattern identification numbers are listed beneath each species designation.

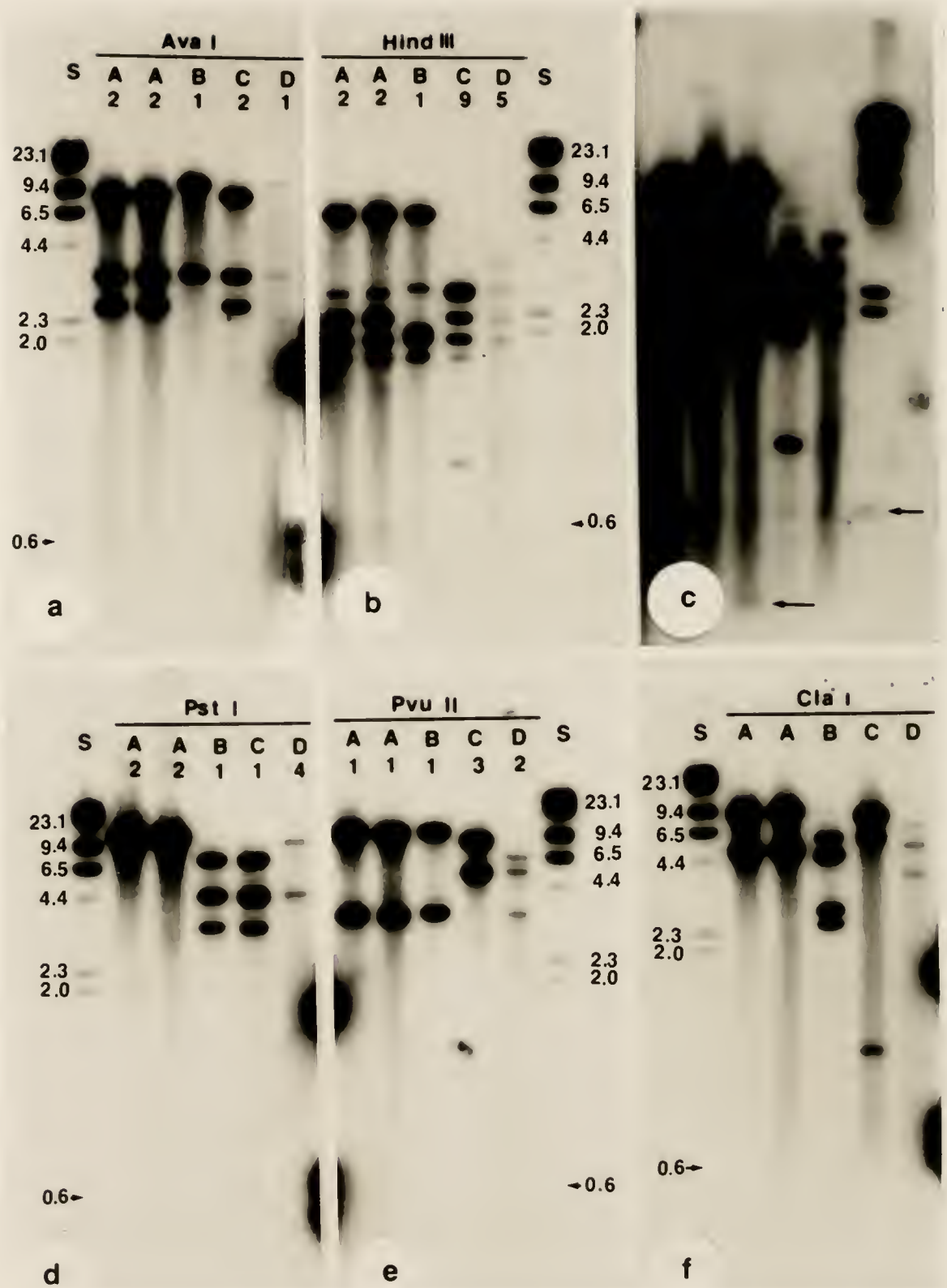
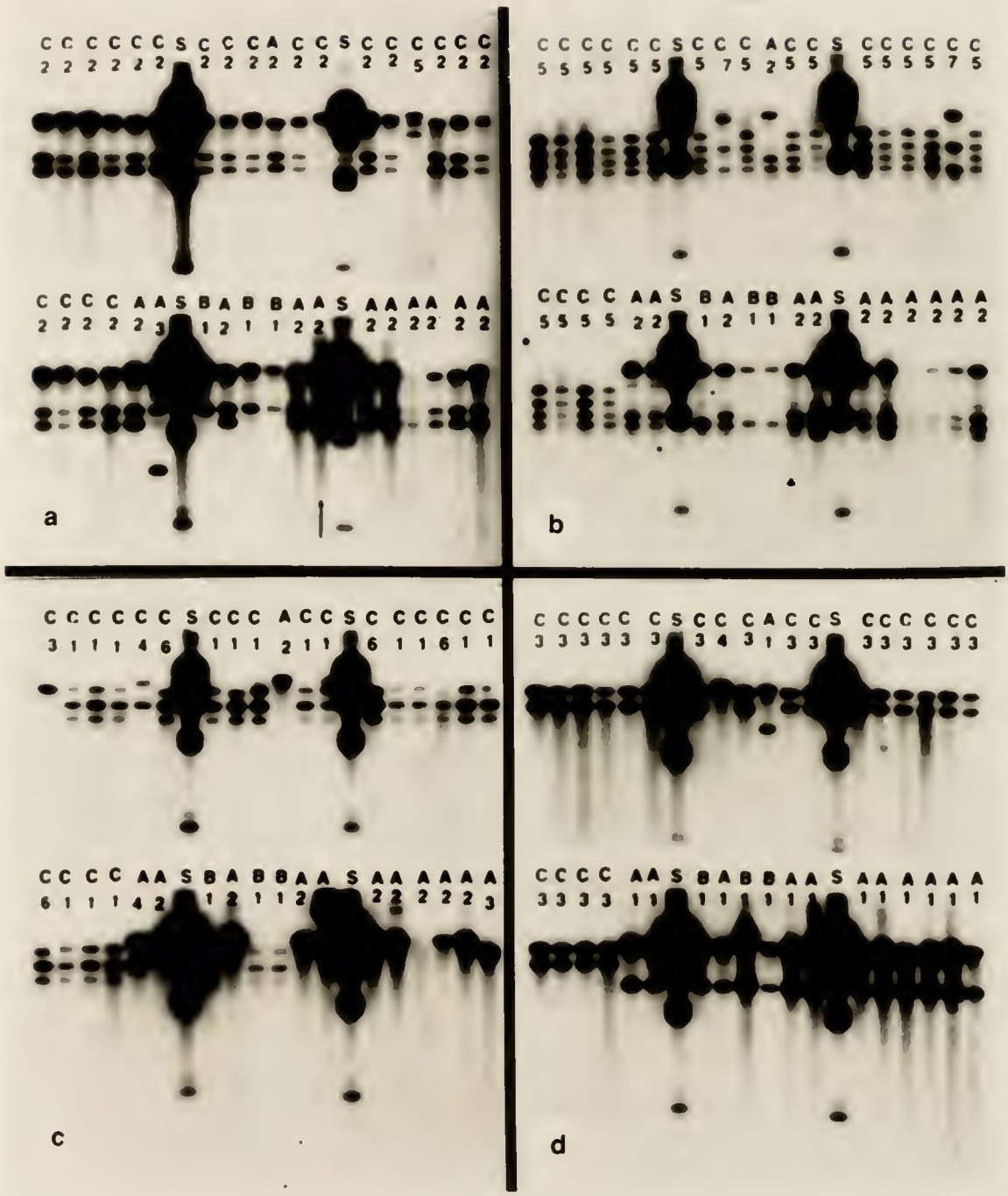

Ava I
S
A 2
A 2
B 1
C 2
D 1
23.1
9.4
6.5
4.4
2.3
2.0
0.6
a
Hind III
A 2
A 2
B 1
C 9
D 5
S
23.1
9.4
6.5
4.4
2.3
2.0
0.6
b
c
Pst I
S
A 2
A 2
B 1
C 1
D 4
23.1
9.4
6.5
4.4
2.3
2.0
0.6
d
Pvu II
A 1
A 1
B 1
C 3
D 2
S
23.1
9.4
6.5
4.4
2.3
2.0
0.6
e
Cla I
S
A
A
B
C
D
23.1
9.4
6.5
4.4
2.3
2.0
0.6
f

Figure 3-3. Representative mitochondrial DNA restriction patterns for field-collected samples of *Anopheles quadrimaculatus* sibling species.

(a) *Ava* I restriction patterns; The arrow on the bottom panel indicates that the aberrant migration of the DNA fragments in this lane is due to distortion of the gel during the Southern blotting procedure and not to differences in molecular weight. (b) *Hind* III restriction patterns; (c) *Pst* I restriction patterns; (d) *Pvu* II restriction patterns.

Note: Letters above the lanes indicate the various species identifications made on the basis of diagnostic isozymes and representative restriction patterns. S=*Hind* III fragments of bacteriophage λ. Pattern identification numbers are listed below each species designation.


a
b
c
d

The *Pst* I cleavage patterns were not useful for species identification. The majority of species B (78%) and species C (85.5%) had the *Pst*-1 pattern (Figures 3-2d and 3-3c). The most common variant in species A was *Pst*-2 which was also present infrequently in the other sibling forms. Species D had the *Pst*-4 restriction pattern which, again, was represented in all the other siblings. It is of interest that this pattern (*Pst*-4) occurred in 43% of the species B mosquitoes collected in Florida but was absent in species B from the Alabama (WHR) and Mississippi (NOX and PIC) populations.

Most of species A (96%) and all the species B mosquitoes examined possessed the same *Pvu* II restriction pattern (*Pvu*-1) (Figures 3-2e and 3-3d). Species D had a unique pattern (*Pvu*-2). Ninety-seven percent of all species C also had a distinctive pattern (*Pvu*-3). There were a few species C individuals, however, that shared a minor pattern (*Pvu*-4) with species A from the Dead Lake population.

None of these four restriction enzymes, when used alone, was capable of discriminating all four sibling species on the basis of their mtDNA cleavage patterns. If data generated by any pair of the three enzymes *Ava* I, *Hind* III or *Pvu* II were combined, however, all siblings could be distinguished. For example, *Hind* III yielded distinctive patterns for species A and B but not for C and D. The latter 2 species could be distinguished by their unique *Pvu* II patterns, however. *Pst* I had little diagnostic value.

The four restriction enzymes used in this study were selected on the basis of preliminary data which indicated that they produced variable patterns among the four sibling species for both the mitochondrial and ribosomal genomes. Even though none of the restriction enzymes used in this population survey produced diagnostic mtDNA

patterns for all of the sibling species, other enzymes such as *Cla* I (Figure 3-2f) had this capability.

Ribosomal DNA analysis

Most of the restriction enzyme digestions that were hybridized to the rDNA probe produced smears of high molecular weight bands along with discrete, smaller DNAs. The indistinct nature of these bands was attributed to length variation between individual copies of the rDNA (probably in the non-transcribed spacer region) and not to poor gel resolution or Southern transfer, since mtDNA bands as large or larger than these were quite sharp.

Species specific rDNA patterns were obtained with *Ava* I, *Hind* III and *Pvu* II (Figures 3-4 and 3-5). For *Ava* I, species A did not have a 2.4 kb fragment that was present in species B (Figure 3-4c). Both species C and D possessed a slightly faster migrating fragment (about 2.3 kb) but species D lacked a 0.7 kb fragment common to the other siblings (Figure 3-4a-c). Species C had a diagnostic 4.8 kb fragment (Figure 3-4b).

The *Hind* III digests also yielded diagnostic rDNA patterns (Figure 3-4d-f). Diagnostic fragments of 1.2 kb in species A, 0.9 kb in species B (Figure 3-4f) and 1.0 kb in species D were obtained (Figure 3-4d). The species C pattern was three-banded with smaller fragments of 1.5 and 1.6 kb (Figure 3-4e). Species A and B also shared a 0.2 kb repeat which was absent in both species C and D (Figure 3-4d and e).

Figure 3-4. Ribosomal DNA patterns of *Anopheles quadrimaculatus* sibling species for *Ava* I and *Hind* III.

(a)-(c) *Ava* I rDNA restriction patterns; (d)-(f) *Hind* III rDNA restriction patterns.

Note: S=*Hind* III fragments from bacteriophage λ. Species designations are listed above each lane. Selected fragment sizes are indicated (kb). Panels (d) and (e) comprise a single gel and illustrate the 0.2 kb repeat common to species A and B but absent in the other siblings. Gels were double-tiered so in panel (d), the 0.2 kb fragment has migrated into the bottom wells (top of panel (e)). The arrow at the top of panel (e) indicates that the 0.2 kb fragment is not present in species D. This small fragment has run off the gel in panel (f).

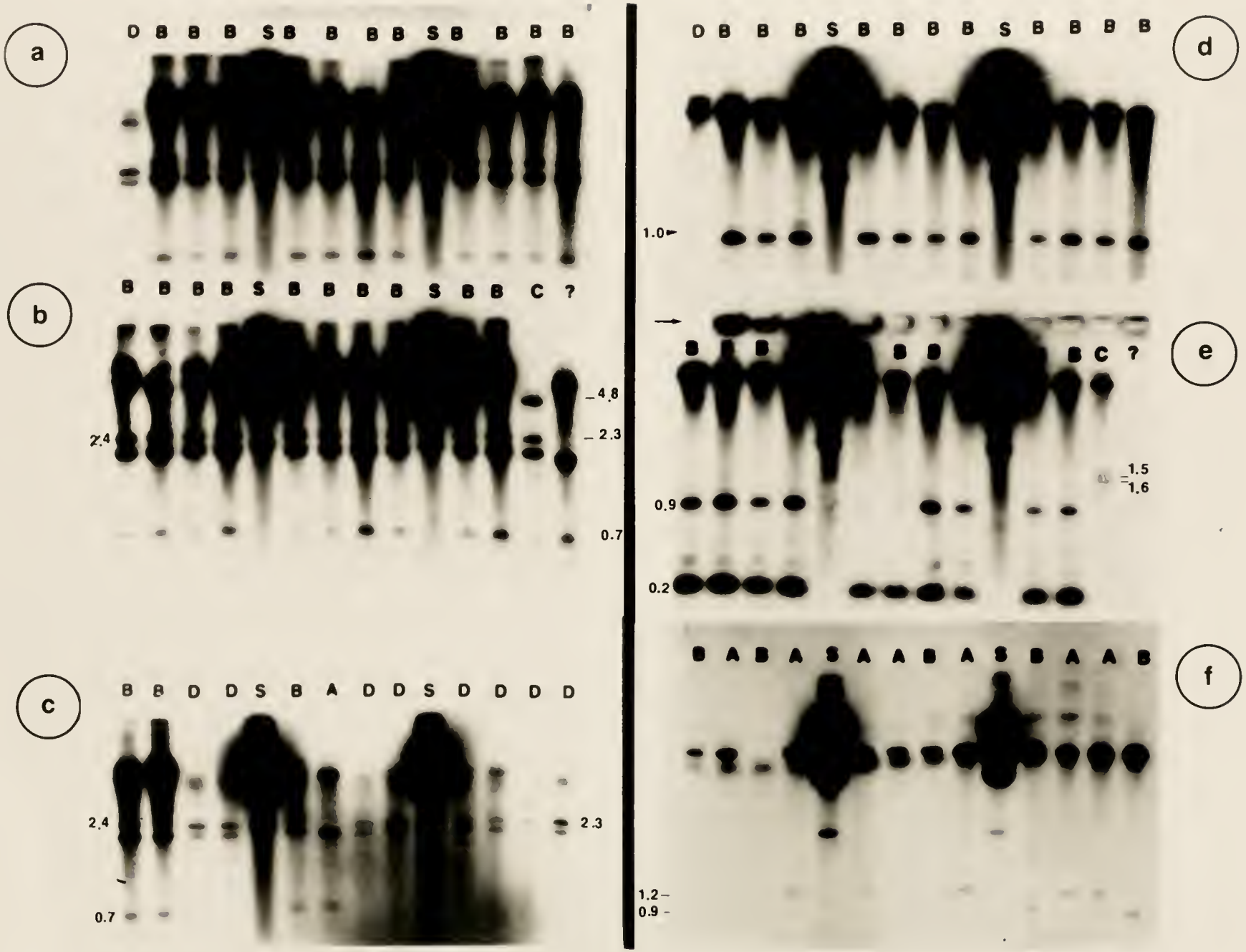
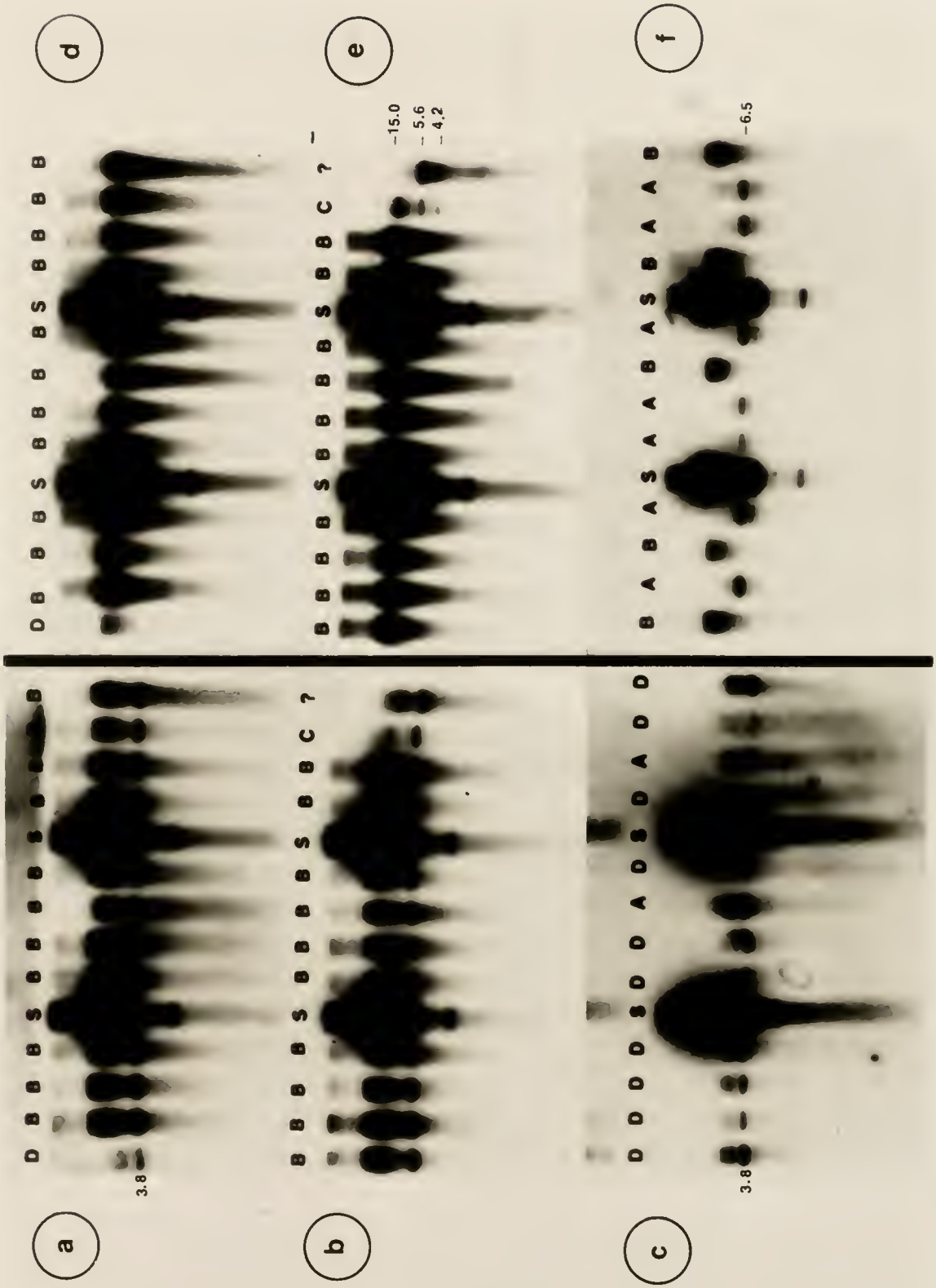

a
D B B B S B B B B S B B B B
b
B B B B S B B B B S B B C ?
2.4
4.8
2.3
0.7
c
B B D D S B A D D S D D D D
2.4
0.7
2.3
d
D B B B S B B B B S B B B B
1.0
e
B B B B C ?
1.5
1.6
0.9
0.2
f
B A B A S A A B A S B A A B
1.2
0.9

Figure 3-5. Ribosomal DNA patterns of *Anopheles quadrimaculatus* sibling species for *Pst* I and *Pvu* II.

(a)-(c) *Pst* I rDNA restriction patterns; (d)-(f) *Pvu* II rDNA restriction patterns.

Note: S=*Hind* III fragments from bacteriophage λ. Species designations are listed above each lane. Selected fragment sizes are indicated (kb).


a
b
c
d
e
f
3.8
3.8
15.0
5.6
4.2
6.5

The *Pst* I rDNA digestion patterns were not particularly illuminating. All four species shared a 3.8 kb fragment (Figure 3-5 a-c).

Restriction patterns for *Pvu* II were specific for species A and C but did not distinguish between species B and D. The latter 2 species contained multiple fragments in the 20-10 kb size range (Figure 3-5d). The species A pattern consisted of a fairly discrete 6.5 kb band (Figure 3-5f) and species C had three fragments (15, 5.6 and 4.2 kb) (Figure 3-5e).

Comparison of Isozyme and DNA Analyses

Species identifications made on the basis of electrophoretic analysis of isozymes and DNA restriction fragment patterns agreed for 99% of the mosquitoes evaluated. Three individuals out of a total of 288 were misidentified by the isozyme procedure.

Two mosquitoes, one from Wheeler Reservoir and one from Noxubee Wild Life Refuge, were identified as species A based on isozymes but these individuals had the mtDNA genotype of species B. Since only the heads of individual mosquitoes were available for protein analysis, a minimal number of diagnostic enzyme loci were examined. Species A and B can be differentiated by using a combination of 2 enzyme loci (*Idh-1* and *Idh-2*) (Lanzaro, 1986). Initially, these were the only two loci employed for identification in populations believed to consist entirely of species A and B. The two mosquitoes in question were heterozygous at the *Idh-1* locus ($Idh\text{-}1^{100}/Idh\text{-}1^{85}$) and homozygous for $Idh\text{-}2^{173}$. Species A is polymorphic at the *Idh-1* locus. In this species, the $Idh\text{-}1^{100}$ allele is the most frequent electromorph (allele frequency = 0.82) but

Idh-1^{85} occurs in about 16% of species A examined from natural populations. Species B, on the other hand, was believed to be fixed for *Idh-1*^{85} (Lanzaro et al., 1990). The *Idh-2*^{173} allele for these mosquitoes was diagnostic for species B (allele frequency = 1.0) but, again, species A is polymorphic at this locus and homozygotes for *Idh-2*^{173} have been encountered (Lanzaro et al., 1990).

Identification of these specimens as species A was primarily indicated by the presence of the *Idh-1*^{100} electromorph. Since the mitochondrial genotype of both mosquitoes was typical of species B, the possibility of interspecific hybridization was considered. Results from the rDNA analysis, however, clearly indicated that these mosquitoes were species B. The rDNA patterns for these individuals are shown in Figure 3-4 panels (b) and (e) (eleventh lane) and Figure 3-4 panel (f) (first lane). It is apparent, therefore, that species B also exhibits a low degree of polymorphism at the *Idh-1* locus.

Another individual, also from Noxubee, was identified as species A on the basis of isozyme analysis but had the mtDNA and rDNA restriction patterns typical of species D. Species D, like species A, is polymorphic at *Idh-1* and *Idh-2* so these loci are incapable of distinguishing these two species (Narang et al., 1989b). Since the Noxubee population was believed to contain only species A and B, diagnostic loci for the other sibling species (such as *Had-3* and *Pgi-1* for species C and *ME-1* and *Got-2* for species D) were not evaluated.

Analysis of Mitochondrial DNA Haplotypes

The simplest measure of polymorphism for a particular segment of DNA (nucleon) is the number of haplotypes (k) observed in the sample (Nei, 1987). This number is dependent on sample size so comparable sample sizes must be used when evaluating the extent of polymorphism present in different populations or species. For the sibling species of the *A. quadrimaculatus* complex, a total of 33 haplotypes were observed for all species combined. Nine haplotypes were found in species A and eight in species B, so these species exhibited about the same amount of polymorphism in their mitochondrial genomes (Table 3-4). Species C was the most polymorphic with a total of 13 haplotypes observed from 2 populations. It was interesting that in the species C populations studied (BBS and LCK), the majority of rare haplotypes were not shared between these populations although they were located only 15 miles apart (Table 3-4). Three haplotypes were noted for species D. The small sample size obtained for this species, however, invalidates any direct comparison to the other three species.

For each species, the mtDNA haplotypes were connected in a "most parsimonious" network (Figure 3-6) which gives the minimum number of mutational steps required to connect all the morphs. The network represents a possible evolutionary pathway for the mtDNA of each species of the *A. quadrimaculatus* complex. These networks, however, do not have directionality and, therefore, do not reveal which one is the most ancestral haplotype. In similar studies, it has been assumed that haplotypes composed exclusively of the major patterns of each enzyme and having the most widespread distribution are the oldest extant haplotypes and are ancestral to the others (Hale and Singh, 1987). For the *A. quadrimaculatus* species

Table 3-4. Distribution of mtDNA haplotpes in natural populations of *Anopheles quadrimaculatus* sibling species.

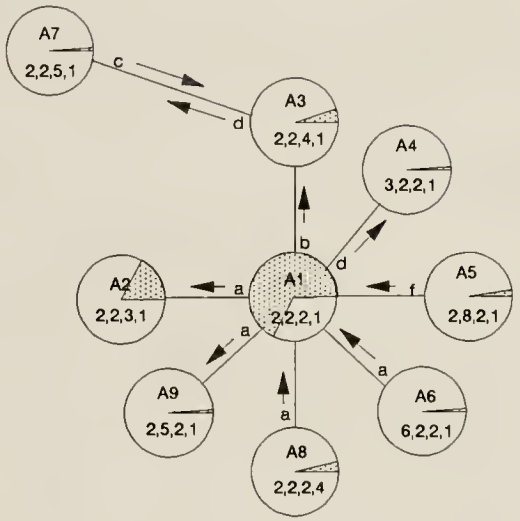
Haplotype[a]		POPULATION							
		LOC	DLK	BBS	LCK	WHR	PIC	NOX	TOTAL
A1	(2,2,2,1)	2	19	1	4	19	2	8	55
A2	(2,2,3,1)	2	4		2	6			14
A3	(2,2,4,1)	2	1			1			4
A4	(3,2,2,1)					1			1
A5	(2,8,2,1)		1					1	2
A6	(6,2,2,1)				1				1
A7	(2,2,5,1)		1						1
A8	(2,2,2,4)		3						3
A9	(2,5,2,1)	1							1
B1	(1,1,1,1)	17	8			21	1	32	79
B2	(1,3,1,1)							2	2
B3	(1,4,1,1)							1	1
B4	(1,1,4,1)	15	4		2				21
B5	(3,1,1,1)	1							1
B6	(1,3,4,1)	1							1
B7	(1,1,2,1)	2							2
B8	(7,1,1,1)	1						1	2
C1	(2,5,1,3)			35	23				58
C2	(2,5,3,3)			5	2				7
C3	(2,5,6,3)				1				1
C4	(4,5,1,3)				1				1
C5	(2,5,1,4)				1				1
C6	(2,6,1,3)				1				1
C7	(1,5,4,3)			1					1
C8	(2,5,4,3)			1					1
C9	(2,7,1,4)			1					1
C10	(5,5,1,3)			1					1
C11	(2,7,1,3)			1					1
C12	(6,5,2,3)			1					1
C13	(2,9,1,3)			1					1
D1	(1,5,4,2)						18	1	19
D2	(1,10,4,2)						1		1
D3	(1,5,2,2)						1		1
Total		44	41	48	38	48	23	46	288

[a] Haplotypes are designated according to species (A1 denotes the major haplotype for species a; B1 for species B, etc.) and by the restriction patterns for Ava 1, Hind III, Pst I and Pvu II, in that order (numbers in parenthesis separated by commas).

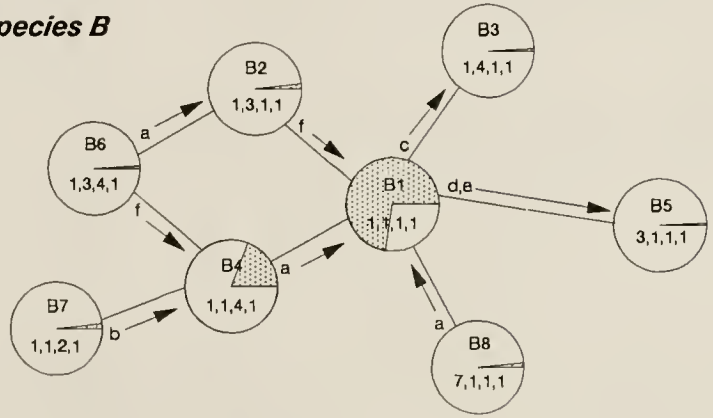
Figure 3-6. Probable phylogenetic relationships among various haplotypes for the four sibling species of the *Anopheles quadrimaculatus* complex.

Note: The diagnostic site(s) between each haplotype is shown. Arrows indicate the direction of the gain of a site but do not necessarily imply direction of evolution. Shaded regions of the "pie diagrams" indicate the proportion of a particular haplotype observed in all populations of that species. The length of the lines connecting the morphs are indicative of the number of restriction site changes required go from one haplotype to the next. Haplotypes with ambiguous origins (i.e. B6 and C9) are connected to more than one parental morph.

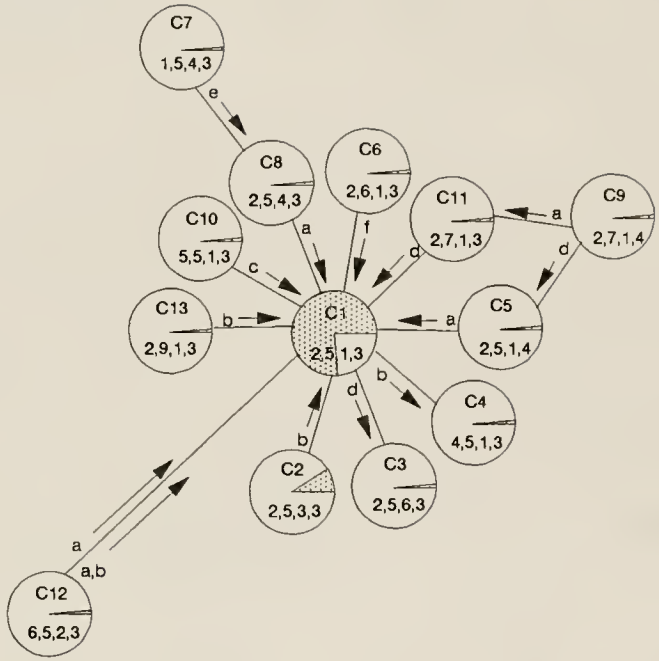
species A



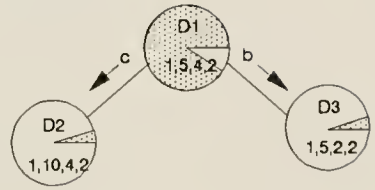
species B



species C



species D



complex, the most common haplotypes have been designated as A1, B1, C1 and D1 for each of the respective sibling species. It is evident from Figure 3-6 that the central positions of these major haplotypes make it necessary to pass through them in order to go from any one pattern to all the others within each species.

Examination of some minor restriction site variants shared between different species, however, did not support the notion that the most common haplotype is always indicative of the ancestral lineage of a species. For example, the *Ava*-3 pattern was present in 1% of the species A and B individuals analyzed. This pattern was created by the presence of one extra site (*Ava*-d) in species A (haplotype A4) and two sites (*Ava*-e and *Ava*-d) in species B (haplotype B5) (Table 3-5). It is parsimonious to assume that a restriction site variant does not arise independently more than once and that the presence of a restriction site is more informative than the absence of a site (because the presence of a site is determined by a sequence consisting of multiple bases but mutation in any one of the bases of the recognition sequence results in the loss of the site). Therefore, the *Ava*-3 pattern present in haplotypes A4 and B5 is most likely to be ancestral for species A and B. Haplotypic frequency, therefore, may merely reflect genotype fitness under current selective pressures or random drift and not the ancestral character state of a species.

Genetic Distance Based on mtDNA Restriction Site Data

Distance calculations for pairwise comparisons for all haplotypes observed in each of the four mosquito species are presented in Table 3-6. Since two classes of

Table 3–5. Presence/absence of varied restriction sites in mtDNA haplotypes of *Anopheles quadrimaculatus* sibling species.

Haplotype[a]	Enzyme: Ava I (a b c d e)	Hind III (a b c d e f)	Pst I (a b c d)	Pvu II (a b)
A1 (2,2,2,1)	+ − + − +	− − − + − +	− − + −	− +
A2 (2,2,3,1)	+ − + − +	− − − + − +	+ − + −	− +
A3 (2,2,4,1)	+ − + − +	− − − + − +	− + + −	− +
A4 (3,2,2,1)	+ − + + +	− − − + − +	− − + −	− +
A5 (2,8,2,1)	+ − + − +	− − − + − −	− − + −	− +
A6 (6,2,2,1)	− − + − +	− − − + − +	− − + −	− +
A7 (2,2,5,1)	+ − + − +	− − − + − +	− + − +	− +
A8 (2,2,2,4)	+ − + − +	− − − + − +	− − + −	− −
A9 (2,5,2,1)	+ − + − +	+ − − + − +	− − + −	− +
B1 (1,1,1,1)	+ − + − −	− − − + + +	+ + + −	− +
B2 (1,3,1,1)	+ − + − −	− − − + + −	+ + + −	− +
B3 (1,4,1,1)	+ − + − −	− − + + + +	+ + + −	− +
B4 (1,1,4,1)	+ − + − −	− − − + + +	− + + −	− +
B5 (3,1,1,1)	+ − + + +	− − − + + +	+ + + −	− +
B6 (1,3,4,1)	+ − + − −	− − − + + −	− + + −	− +
B7 (1,1,2,1)	+ − + − −	− − − + + +	− − + −	− +
B8 (7,1,1,1)	+ − − − −	− − − + + +	+ + + −	− +
C1 (2,5,1,3)	+ − + − +	+ − − + − +	+ + + −	+ −
C2 (2,5,3,3)	+ − + − +	+ − − + − +	+ − + −	+ −
C3 (2,5,6,3)	+ − + − +	+ − − + − +	+ + + +	+ −
C4 (4,5,1,3)	+ + + − +	+ − − + − +	+ + + −	+ −
C5 (2,5,1,4)	+ − + − +	+ − − + − +	+ + + −	− −
C6 (2,6,1,3)	+ − + − +	+ − − + − −	+ + + −	+ −
C7 (1,5,4,3)	+ − + − −	+ − − + − +	− + + −	+ −
C8 (2,5,4,3)	+ − + − +	+ − − + − +	− + + −	+ −
C9 (2,7,1,4)	+ − + − +	+ − − − − +	+ + + −	− −
C10 (5,5,1,3)	+ − − − +	+ − − + − +	+ + + −	+ −
C11 (2,7,1,3)	+ − + − +	+ − − − − +	+ + + −	+ −
C12 (6,5,2,3)	− − + − +	+ − − + − +	− − + −	+ −
C13 (2,9,1,3)	+ − + − +	+ + − + − +	+ + + −	+ −
D1 (1,5,4,2)	+ − + − −	+ − − + − +	− + + −	+ +
D2 (1,10,4,2)	+ − + − −	+ − + + − +	− + + −	+ +
D3 (1,5,2,2)	+ − + − −	+ − − + − +	− − + −	+ +

[a] Haplotypes are designated according to species (A1 denotes the major haplotype for species A; B1 for species B, etc.) and by the restriction patterns for Ava 1, Hind III, Pst I and Pvu II, in that order (numbers in parenthesis separated by commas).

Note: A plus sign (+) denotes presence of the site; a minus sign (−) denotes absence of the site.

Table 3-6. Distance matrix for mtDNA haplotypes observed in all sibling species of the *Anopheles quadrimaculatus* complex.

	A1	A2	A3	A4	A5	A6	A7	A8	A9	B1	B2	B3	B4	B5	B6	B7	B8	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	D1	D2	D3
A1		0.009	0.009	0.000	0.010	0.000	0.029	0.010	0.009	0.026	0.037	0.033	0.018	0.026	0.029	0.009	0.026	0.045	0.037	0.053	0.045	0.037	0.059	0.037	0.037	0.051	0.045	0.059	0.029	0.053	0.026	0.033	0.018
A2	0.000		0.018	0.009	0.020	0.009	0.037	0.020	0.018	0.016	0.026	0.023	0.026	0.016	0.037	0.018	0.016	0.033	0.026	0.041	0.033	0.026	0.045	0.045	0.045	0.037	0.033	0.045	0.037	0.041	0.033	0.041	0.026
A3	0.000	0.000		0.009	0.020	0.009	0.018	0.020	0.018	0.016	0.026	0.023	0.008	0.016	0.018	0.018	0.016	0.033	0.045	0.041	0.033	0.026	0.045	0.026	0.026	0.037	0.033	0.045	0.037	0.041	0.016	0.023	0.026
A4	0.031	0.031	0.031		0.010	0.000	0.029	0.010	0.009	0.026	0.037	0.033	0.018	0.026	0.029	0.009	0.026	0.045	0.037	0.053	0.045	0.037	0.059	0.037	0.037	0.051	0.045	0.059	0.029	0.053	0.026	0.033	0.018
A5	0.000	0.000	0.000	0.031		0.010	0.042	0.022	0.020	0.037	0.029	0.045	0.029	0.037	0.020	0.020	0.037	0.059	0.051	0.068	0.059	0.051	0.051	0.051	0.051	0.068	0.059	0.077	0.042	0.068	0.037	0.045	0.029
A6	0.045	0.045	0.045	0.081	0.045		0.029	0.010	0.009	0.026	0.037	0.033	0.018	0.026	0.029	0.009	0.026	0.045	0.037	0.053	0.045	0.037	0.059	0.037	0.037	0.051	0.045	0.059	0.029	0.053	0.026	0.033	0.018
A7	0.000	0.000	0.000	0.031	0.000	0.045		0.042	0.037	0.033	0.045	0.041	0.026	0.033	0.037	0.037	0.033	0.053	0.068	0.041	0.053	0.045	0.068	0.045	0.045	0.059	0.053	0.068	0.059	0.060	0.033	0.041	0.045
A8	0.000	0.000	0.000	0.031	0.000	0.045	0.000		0.020	0.037	0.051	0.045	0.029	0.037	0.042	0.020	0.037	0.037	0.029	0.045	0.037	0.029	0.051	0.029	0.029	0.042	0.037	0.051	0.020	0.045	0.037	0.045	0.029
A9	0.000	0.000	0.000	0.031	0.000	0.045	0.000	0.000		0.033	0.045	0.041	0.026	0.033	0.037	0.018	0.033	0.033	0.026	0.041	0.033	0.026	0.045	0.026	0.026	0.037	0.033	0.045	0.018	0.041	0.016	0.023	0.008
B1	0.045	0.045	0.045	0.081	0.045	0.139	0.045	0.045	0.045		0.007	0.007	0.007	0.000	0.016	0.016	0.000	0.030	0.041	0.037	0.030	0.023	0.041	0.041	0.041	0.033	0.030	0.041	0.053	0.037	0.030	0.037	0.041
B2	0.045	0.045	0.045	0.081	0.045	0.139	0.045	0.045	0.045	0.000		0.015	0.016	0.007	0.008	0.026	0.007	0.041	0.053	0.048	0.041	0.033	0.033	0.053	0.053	0.045	0.041	0.053	0.068	0.048	0.041	0.048	0.053
B3	0.045	0.045	0.045	0.081	0.045	0.139	0.045	0.045	0.045	0.000	0.000		0.015	0.007	0.023	0.023	0.007	0.037	0.048	0.044	0.037	0.030	0.048	0.048	0.048	0.041	0.037	0.048	0.060	0.044	0.037	0.028	0.048
B4	0.045	0.045	0.045	0.081	0.045	0.139	0.045	0.045	0.045	0.000	0.000	0.000		0.007	0.008	0.008	0.007	0.041	0.053	0.048	0.041	0.033	0.053	0.033	0.033	0.045	0.041	0.053	0.045	0.048	0.023	0.030	0.033
B5	0.031	0.031	0.031	0.000	0.031	0.081	0.031	0.031	0.031	0.081	0.081	0.081	0.081		0.016	0.016	0.000	0.030	0.041	0.037	0.030	0.023	0.041	0.041	0.041	0.033	0.030	0.041	0.053	0.037	0.030	0.037	0.041
B6	0.045	0.045	0.045	0.081	0.045	0.139	0.045	0.045	0.045	0.000	0.000	0.000	0.000	0.081		0.018	0.016	0.053	0.068	0.060	0.053	0.045	0.045	0.045	0.045	0.059	0.053	0.068	0.059	0.060	0.033	0.041	0.045
B7	0.045	0.045	0.045	0.081	0.045	0.139	0.045	0.045	0.045	0.000	0.000	0.000	0.000	0.081	0.000		0.016	0.053	0.045	0.060	0.053	0.045	0.068	0.045	0.045	0.059	0.053	0.068	0.037	0.060	0.033	0.041	0.026
B8	0.139	0.139	0.139	0.183	0.139	0.200	0.139	0.139	0.139	0.081	0.081	0.081	0.081	0.183	0.081	0.081		0.030	0.041	0.037	0.030	0.023	0.041	0.041	0.041	0.033	0.030	0.041	0.053	0.037	0.030	0.037	0.041
C1	0.000	0.000	0.000	0.031	0.000	0.045	0.000	0.000	0.000	0.045	0.045	0.045	0.045	0.031	0.045	0.045	0.139		0.007	0.007	0.000	0.007	0.007	0.007	0.007	0.016	0.000	0.007	0.016	0.007	0.015	0.021	0.023
C2	0.000	0.000	0.000	0.031	0.000	0.045	0.000	0.000	0.000	0.045	0.045	0.045	0.045	0.031	0.045	0.045	0.139	0.000		0.015	0.007	0.016	0.016	0.016	0.016	0.026	0.007	0.016	0.008	0.015	0.023	0.030	0.016
C3	0.000	0.000	0.000	0.031	0.000	0.045	0.000	0.000	0.000	0.045	0.045	0.045	0.045	0.031	0.045	0.045	0.139	0.000	0.000		0.007	0.015	0.015	0.015	0.015	0.023	0.007	0.015	0.023	0.013	0.021	0.028	0.030
C4	0.031	0.031	0.031	0.058	0.031	0.081	0.031	0.031	0.031	0.081	0.081	0.081	0.081	0.058	0.081	0.081	0.183	0.031	0.031	0.031		0.007	0.007	0.007	0.007	0.016	0.000	0.007	0.016	0.007	0.015	0.021	0.023
C5	0.000	0.000	0.000	0.031	0.000	0.045	0.000	0.000	0.000	0.045	0.045	0.045	0.045	0.031	0.045	0.045	0.139	0.000	0.000	0.000	0.031		0.016	0.016	0.016	0.008	0.007	0.016	0.026	0.015	0.023	0.030	0.033
C6	0.000	0.000	0.000	0.031	0.000	0.045	0.000	0.000	0.000	0.045	0.045	0.045	0.045	0.031	0.045	0.045	0.139	0.000	0.000	0.000	0.031	0.000		0.016	0.016	0.026	0.007	0.016	0.026	0.015	0.023	0.030	0.033
C7	0.045	0.045	0.045	0.081	0.045	0.139	0.045	0.045	0.045	0.000	0.000	0.000	0.000	0.081	0.000	0.000	0.081	0.045	0.045	0.045	0.081	0.045	0.045		0.000	0.026	0.007	0.016	0.008	0.015	0.007	0.015	0.016
C8	0.000	0.000	0.000	0.031	0.000	0.045	0.000	0.000	0.000	0.045	0.045	0.045	0.045	0.031	0.045	0.045	0.139	0.000	0.000	0.000	0.031	0.000	0.000	0.045		0.026	0.007	0.016	0.008	0.015	0.007	0.015	0.016
C9	0.000	0.000	0.000	0.031	0.000	0.045	0.000	0.000	0.000	0.045	0.045	0.045	0.045	0.031	0.045	0.045	0.139	0.000	0.000	0.000	0.031	0.000	0.000	0.045	0.000		0.016	0.008	0.037	0.023	0.033	0.041	0.045
C10	0.045	0.045	0.045	0.081	0.045	0.139	0.045	0.045	0.045	0.139	0.139	0.139	0.139	0.081	0.139	0.139	0.081	0.045	0.045	0.045	0.081	0.045	0.045	0.139	0.045	0.045		0.007	0.016	0.007	0.015	0.021	0.023
C11	0.000	0.000	0.000	0.031	0.000	0.045	0.000	0.000	0.000	0.045	0.045	0.045	0.045	0.031	0.045	0.045	0.139	0.000	0.000	0.000	0.031	0.000	0.000	0.045	0.000	0.000	0.045		0.026	0.015	0.023	0.030	0.033
C12	0.045	0.045	0.045	0.081	0.045	0.000	0.045	0.045	0.045	0.139	0.139	0.139	0.139	0.081	0.139	0.139	0.200	0.045	0.045	0.045	0.081	0.045	0.045	0.139	0.045	0.045	0.139	0.045		0.023	0.016	0.023	0.008
C13	0.000	0.000	0.000	0.031	0.000	0.045	0.000	0.000	0.000	0.045	0.045	0.045	0.045	0.031	0.045	0.045	0.139	0.000	0.000	0.000	0.031	0.000	0.000	0.045	0.000	0.000	0.045	0.000	0.045		0.021	0.028	0.030
D1	0.045	0.045	0.045	0.081	0.045	0.139	0.045	0.045	0.045	0.000	0.000	0.000	0.000	0.081	0.000	0.000	0.081	0.045	0.045	0.045	0.081	0.045	0.045	0.000	0.045	0.045	0.139	0.045	0.139	0.045		0.007	0.007
D2	0.045	0.045	0.045	0.081	0.045	0.139	0.045	0.045	0.045	0.000	0.000	0.000	0.000	0.081	0.000	0.000	0.081	0.045	0.045	0.045	0.081	0.045	0.045	0.000	0.045	0.045	0.139	0.045	0.139	0.045	0.000		0.015
D3	0.045	0.045	0.045	0.081	0.045	0.139	0.045	0.045	0.045	0.000	0.000	0.000	0.000	0.081	0.000	0.000	0.081	0.045	0.045	0.045	0.081	0.045	0.045	0.000	0.045	0.045	0.139	0.045	0.139	0.045	0.000	0.000	

Note: Distances are expressed as the mean number of nucleotide substitutions per nucleotide site (δ). Values above the diagonal are for Hind III, Pst I and Pvu II combined (r*=6),and distances calculated for Ava I (r=5.33) are listed below the diagonal.

* r = number of bases in the restriction enzyme recognition sequence; r value for Ava I was calculated according to Nei, 1987.

restriction enzymes were used (*Hind* III, *Pst* I and *Pvu* II each recognize one unique six-base sequence while *Ava* I recognizes four different six-base sequences), distances were reported separately for each enzyme class.

Pooled distance calculations based on all mtDNA restriction site variants observed within each species for both enzyme classes are reported in Table 3-7. Distance values are presented for each separate enzyme class as well as combined weighted distances for both classes. The pooled distance values were weighted by the proportion of the total number of sites contributed by each enzyme class across all species (1:5 weighting).

The genetic distances calculated from restriction site data indicated that *A. quadrimaculatus* species A is more closely related to species B than to species C or D. Species C and D also cluster and an unrooted tree based on these data is presented in Figure 3-7.

Discussion

The data presented in this study confirm that restriction site polymorphisms unique to the mitochondrial and nuclear (rDNA) genomes of members of the *A. quadrimaculatus* complex can be used for identification of individual mosquitoes to species. The DNA-based techniques were slightly more labor-intensive than isozyme analysis but the molecular methods proved to be more reliable for correct species assignment than protein electrophoresis, especially when a small number of diagnostic isozymes was evaluated.

Table 3-7. Pooled interspecific genetic distances for all haplotypes and including both restriction enzyme classes.

Weight[a] A:HPV	Species	Genetic distance (δ) A	B	C	D
1:5	A	0.00485 (0.00080)[b]	0.02709 (0.00470)	0.03574 (0.00652)	0.02993 (0.00521)
	B		0.00352 (0.00059)	0.03682 (0.00661)	0.02548 (0.00449)
	C			0.00347 (0.00058)	0.02117 (0.00364)
	D				0.00108 (0.00018)
1:0	A	0.00171 (0.00030)	0.04484 (0.00799)	0.00290 (0.00051)	0.04336 (0.00760)
	B		0.00413 (0.00075)	0.04554 (0.00820)	0.00209 (0.00038)
	C			0.00407 (0.00071)	0.04409 (0.00783)
	D				0.00000 (0.00000)
0:1	A	0.00547 (0.00090)	0.02354 (0.00405)	0.04231 (0.00773)	0.02725 (0.00474)
	B		0.00340 (0.00056)	0.03508 (0.00629)	0.03016 (0.00531)
	C			0.00335 (0.00055)	0.01659 (0.00280)
	D				0.00129 (0.00021)

[a] Weighting factor for *Ava* I (A) and for *Hind* III, *Pst* 1 and *Pvu* II combined (HPV).

[b] Variance associated with the estimation of δ is given in parentheses.

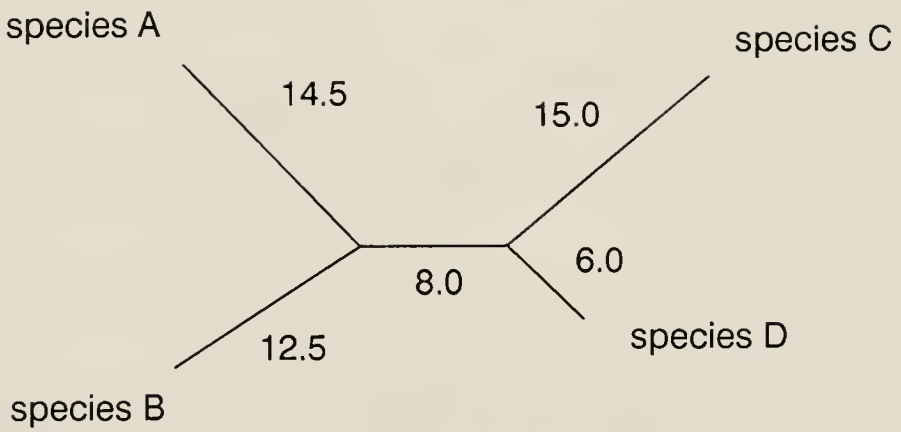


Figure 3-7. Unrooted phylogenetic tree for the four sibling species of the *Anopheles quadrimaculatus* complex.

Note: This tree was based on pooled distance values derived from restriction site analysis of the mitochondrial genomes of all sibling species. Branch lengths denote the number of substitutions per kilobase pair.

The four siblings species of the *A. quadrimaculatus* complex are sympatric and in some populations adults of all species can be collected from the same tree hole. There has been no genetic evidence (from extensive cytological studies or from surveys designed to detect hybrid males based on abnormal testes morphology) for interspecific hybridization between species A and B in the wild (Lanzaro, 1986). One presumed A/C hybrid, however, has been detected by DNA hybridization to cloned species-specific repetitive sequences (Cockburn, 1990). If interspecific hybridization occurs between species A and C, it does so at an extremely low level since the supposed hybrid individual comprised only 0.3% of the mosquitoes analyzed (Cockburn, 1990).

The present study represents the first attempt at a systematic correlation of the cytoplasmic and nuclear genomes of individual mosquitoes from mixed populations of *A. quadrimaculatus* sibling species. The fact that DNA analysis found no evidence for interspecific hybridization confirms previous observations and indicates that strong prezygotic isolating mechanisms must be operating to preclude matings among the various sibling forms in their natural environment.

Recently, it has been noted that heteroplasmy, in which two mtDNA variants occur abundantly in an individual, occurs fairly frequently in both vertebrate and invertebrate taxa (Hale and Singh, 1986; Harrison et al., 1985; Hauswirth and Laipis, 1982; Solignac et al, 1983). A few of the mtDNA restriction patterns generated by *Ava* I, at first glance, indicated that some individuals (both species A and species C) might be heteroplasmic for the *Ava*-1 and *Ava*-2 patterns (Figure 3-3a, most lanes of top panel). Further analysis revealed, however, that these mixed patterns resulted from recalcitrant digestion of the *Ava*-e site and not heteroplasmy.

Variation in the length of the A+T rich region of the mitochondrial genome is also common in a variety of different animals (Brown, 1987). Since this region does not code for a gene product, deletions and/or insertions are readily maintained without apparent harm to the organism. Polymorphism in the length of the A+T-rich region, as evidenced by slight differences in the mobility of the 2.7 kb *Hind* III fragment, was noted for both species A and species B of the *A. quadrimaculatus* complex.

Several genetic studies have shown that species A and B are more related to each other than either is to the other two sibling forms. This conclusion has been supported by cytological, isozyme and hybridization studies (Kaiser et al., 1989; Narang et al., 1989a and 1989b). Phenograms of genetic relationships based on electrophoretic data have not yet included species D, but cytological analyses indicate that species C and D are probably more closely related to each other than to species A or B (Kaiser et al., 1989; Narang et al. 1989b). The genetic distances calculated from the mtDNA restriction site data and the clustering of taxonomic units in the unrooted tree generated from these data generally agreed with the interspecific relationships described above. However, some of the genetic relationships estimated from restriction site data were contrary to those derived from hybridization experiments. For example, hybridization studies have shown that: (1) species C is more related to species B than to species A, and; (2) species A is more related to species C than to species D (Kaiser et al., 1989; Narang et al.,1989b). The genetic distances reported in this study suggest that species A is closer to species D than to C, and that species A is more related to species C than species B is. The distance data also suggest that species B is slightly closer to species D than to

species A. Unfortunately, hybridizations between species B and D have not yet been performed.

The validity of the genetic distance values derived for members of the *A. quadrimaculatus* complex was influenced by two factors. First of all, the interspecific distances were very small. For example, the distances calculated for the mosquito siblings ranged from about 2-4% as compared to the 5-10% nucleotide divergence reported for 8 species of the *D. melanogaster* subgroup (Solignac et al, 1986). Secondly, the variances associated with the estimation of δ were fairly large, averaging about 16% the value of δ. The variance of δ is large when the average number of restriction sites surveyed is small (Nei and Li, 1979). Therefore, the reliability of the estimate is increased if many different restriction enzymes are used. Since the primary objective of this study was to develop a simple and rapid means of identifying large numbers of mosquito species, a limited number of restriction enzymes was employed. Nevertheless, many of the interspecific relationships derived from these distance calculations were biologically meaningful and were in agreement with relationships deduced from other genetic methods.

There are several methods used for reconstructing phylogenetic trees from molecular data. Distance matrix methods, such as the neighbor-joining method used in this study, employ genetic distances computed for all pairs of species and a phylogenetic tree is constructed by considering the relationships among the distance values. Maximum parsimony methods, on the other hand, deduce the nucleotide sequences of ancestral species from those of extant species and a tree is produced by minimizing the number of evolutionary changes for the entire tree. Trees constructed by parsimony

methods rely on the existence of "informative" restriction or nucleotide sites (Nei, 1987). A site is informative only if there are at least two kinds of sites, each represented at least two times. For example, in the *A. quadrimaculatus* complex a restriction site is informative only if it is shared between two species and is absent in the other two siblings. There are a total of five informative restriction sites produced by the restriction enzymes used in this study. Three of these sites (*Pvu*-a, *Ava*-d and *Hind*-a) are shared by species A and B and species C and D, respectively. Conversely, the *Ava*-e site is present in the majority of species A and C and the rare *Hind*-c site is shared by species B and D. Three informative sites, therefore, cluster species A with B and C with D while the other two sites indicate that A/C and B/D is the correct relationship. In order for maximum parsimony methods to be used to discern the relationships among members of the *A. quadrimaculatus* species complex, a more thorough study utilizing many different restriction enzymes is needed.

The value of using mtDNA restriction site data in studying closely related species has been highly touted because of the apparent rapid evolution of mtDNA in vertebrates (Brown et al. 1979; Avise et al., 1979). It has been estimated that the mtDNA is evolving about ten times more rapidly than the nuclear genome in mammals (Brown, 1985). The nuclear genome of insects, however, seems to be evolving at a much higher rate than vertebrate nuclear DNAs. For example, in a study of *D. melanogaster* and *D. simulans*, it has been suggested that mtDNA and allozyme loci evolve at similar rates (Shah and Langley, 1979). Comparisons of the reassociation kinetics of mtDNA and single copy nuclear DNAs in *D. melanogaster* have confirmed

this observation (Powell et al., 1986). If further studies on insects substantiate these findings, the implication is that mtDNA restriction analysis may not prove to be as useful for studying closely related insect species as it has for vertebrates.

CHAPTER IV

CONCLUDING REMARKS

The major thrust of the work reported in this study was twofold. First, the mitochondrial genome of *Anopheles quadrimaculatus,* species A was cloned and characterized. Secondly, these mtDNA clones, along with a previously isolated rDNA clone, were used as radioactive probes for identification of species-specific DNA restriction patterns for individual mosquitoes from natural populations containing more than one sibling species of *A. quadrimaculatus.*

The sequence content and organization of both the mitochondrial genome and nuclear genes that code for ribosomal RNAs have been highly conserved over the course of evolution. The DNA probes that were developed for *A. quadrimaculatus,* therefore, can be used as a general tool for the investigation of genetic differences in any mosquito or even more distantly-related species.

Anopheline mosquitoes belonging to the subgenus Nyssorhynchus are extremely difficult to differentiate taxonomically. Some species of this subgenus are important vectors of malaria in South and Central America. Preliminary results indicate that the mtDNA probes developed as a result of this study can be used to differentiate many of these species (e.g. *A. nuneztovari, A. triannulatus, A. oswaldoi, A. marajoara* and *A. albimanus*). The mtDNA probes will be used for further study of the genetic

relationships among these important mosquitoes. Hopefully, knowledge gained by careful study of the genetics, vector capacity and ecology of anopheline mosquitoes will ultimately contribute to a reduction in the incidence of malaria in the third-world.

LITERATURE CITED

Avise, J.C., C. Giblin-Davidson, J. Laerm, J.C. Patton, and R.A. Lansman. Mitochondrial DNA clones and matriarchal phylogeny within and among geographic populations of the pocket gopher, *Geomys pinetis*. Proc. Natl. Acad. Sci. USA 76: 6694-6698. (1979).

Avise, J.C., R.A. Lansman, and R.O. Shade. The use of restriction endonucleases to measure mitochondrial DNA sequence relatedness in natural populations. I. Population structure and evolution in the genus *Peromyscus*. Genetics 92: 279-295. (1979).

Ayala, F.J., and J.R. Powell. Allozymes as diagnostic characters in sibling species of *Drosophila*. Proc. Natl. Acad. Sci. USA 69: 1094-1096. (1972).

Berger, S.L., and A.R. Kimmel. Guide to Molecular Cloning Techniques. Methods in Enzymology. Academic Press, San Diego. Vol. 152. (1987).

Birky, C.W., Jr, T. Maruyama, and P. Fuerst. An approach to population and evolutionary genetic theory for genes in mitochondria and chloroplasts and some results. Genetics 103: 513-527. (1983).

Brown, W.M. The mitochondrial genome of animals. In: Molecular Evolutionary Genetics (R.J. MacIntyre, ed.). Plenum, New York. pp. 95-130. (1985).

Brown, W.M., M. George, Jr., and A.C. Wilson. Rapid evolution of animal mitochondrial DNA. Proc. Natl. Acad. Sci. USA 76: 1967-1971. (1979).

Chomyn, A., P. Mariottini, M.W.J. Cleeter, C.I. Ragan, A. Matsuno-Yagi, Y. Hatefi, R.F. Doolittle, and G. Attardi. Six unidentified reading frames of human mitochondrial DNA encode components of the respiratory-chain NADH dehydrogenase. Nature 314: 592-597. (1985).

Clary, D.O., and D.R. Wolstenholme. Genes for cytochrome c oxidase subunit I, URF2 and three tRNAs in *Drosophila* mitochondrial DNA. Nucleic Acids Res. 11: 6859-6872. (1983).

Clary, D.O., and D.R. Wolstenholme. The mitochondrial DNA molecule of *Drosophila yakuba*: nucleotide sequence, gene organization, and genetic code. J. Mol. Evol. 22: 252-271. (1985).

Cockburn, A.F. A simple and rapid technique for identification of large numbers of individual mosquitoes using DNA hybridization. Arch. Insect Biochem. Physiol. In press. (1990).

Cockburn, A.F., and S.E. Mitchell. Repetitive DNA interspersion patterns in Diptera. Arch. Insect Biochem. Physiol. 10: 105-113. (1989).

Cockburn, A.F., and J.A. Seawright. Techniques for mitochondrial and ribosomal DNA analysis of anopheline mosquitoes. J. Am. Mosq. Control Assoc. 3: 261-265. (1988).

Darsie, R.F., Jr., and R.A. Ward. Identification and geographical distribution of the mosquitoes of North America, north of Mexico. Mosq. Syst. Suppl. 1: 1-313. (1981).

Dawid, I. Evolution of mitochondrial DNA sequences in *Xenopus*. Dev. Biol. 29: 139-151. (1972).

Devereaux, J., P. Haeberli, and O. Smithies. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. 12: 387-394. (1984).

Dubin, D.T., C.C. HsuChen, and L.E. Tillotson. Mosquito mitochondrial transfer RNAs for valine, glycine, and glutamate: RNA and gene sequences and vicinal genome organization. Curr. Genet. 10: 701-707. (1986).

Erlich, D. Illegitimate recombination in bacteria. In: Mobile DNA (D.E. Berg and M.M. Howe, eds.). ASM Publications, Washington, DC. pp. 799-824. (1989).

Garesse, R. *Drosophila melanogaster* mitochondrial DNA: Gene organization and evolutionary considerations. Genetics 118: 649-663. (1988).

Gyllensten, U., D. Wharton, and A.C. Wilson. Maternal inheritance of mitochondrial DNA during backcrossing of two species of mice. J. Hered. 76: 321-324. (1985).

Hale, L.H., and R.S. Singh. Extensive variation and heteroplasmy in size of mitochondiral DNA among geographic populations of *Drosophila melanogaster*. Proc. Natl. Acal. Sci. USA 83: 8813-8817. (1986).

Hale, L.H., and R.S. Singh. Mitochondrial DNA variation and genetic structure in populations of *Drosophila melanogaster*. Mol. Biol. Evol. 4: 622-637. (1987).

Harrison, R.G., D.M. Rand, and W.C. Wheeler. Mitochondrial DNA size variation within individual crickets. Science 228: 1446-1448. (1985).

Hauswirth, W.W., and P.J. Laipis. Mitochondrial DNA polymorphism in a maternal lineage of Holstein cows. Proc. Natl. Acad. Sci. USA 79: 4686-4690. (1982).

HsuChen, C.C., R.M. Kotin, and D.T. Dubin. Sequences of the coding and flanking regions of the large ribosomal subunit RNA of mosquito mitochondria. Nucleic Acids Res. 2: 7771-7785. (1984).

Hutchinson, C. A., III, J.E. Newbold, S.S. Potter, and M.H. Edgell. Maternal inheritance of mammalian mitochondrial DNA. Nature 251: 536-538. (1974).

Kaiser, P.E., S.E. Mitchell, G.C. Lanzaro, and J.A. Seawright. Hybridization of laboratory strains of sibling species A and B of *Anopheles quadrimaculatus.* J. Am. Mosq. Control Assoc. 4: 34-38. (1988a).

Kaiser, P.E., S.K. Narang, J.A. Seawright, and D. Kline. A new member of the *Anopheles quadrimaculatus* complex, species C. J. Amer. Mosq. Control Assoc. 4: 494-499. (1988b).

Kitzmiller, J.B. Mosquito genetics and cytogenetics. Separata da Rev. Bras. de Malariol. e D. Trop. 5: 285-359. (1953).

Kitzmiller, J.B., G. Grizzi, and R.H. Baker. Evolution and speciation within the *maculipennis* complex of the genus *Anopheles.* In: Genetics of Insect Vectors of Disease (J.W. Wright and R. Pal, eds.). Elsevier Publ. Co., Amsterdam. pp. 151-210. (1967).

Lansman, R.A., J.C. Avise, and M.D. Huettel. Critical experimental test of the possibility of "paternal leakage" of mitochondrial DNA. Proc. Natl. Acad. Sci. USA 80: 1969-1971. (1983).

Lansman, R.A., R.O. Shade, J.F. Shapira, and J.C. Avise. The use of restriction endonucleases to measure mitochondrial DNA sequence relatedness in natural populations. III. Techniques and potential applications. J. Mol. Evol. 17: 214-226. (1981).

Lanzaro, G. Use of enzyme polymorphism and hybridization crosses to identify sibling species of the mosquito, *Anopheles quadrimaculatus* (Say). Ph.D. dissertation. University of Florida, Gainesville, Florida. 92 pp. (1986).

Lanzaro, G.C., S.K. Narang, S.E. Mitchell, P.E. Kaiser, and J.A. Seawright. Hybrid male sterility in crosses between field and laboratory strains of *Anopheles quadrimaculatus* (Say) (Diptera: Culicidae). J. Med. Entomol. 25: 248-255. (1988).

Lanzaro, G.C., S.K. Narang, and J.A. Seawright. Speciation in an anopheline (Diptera:Culicidae) mosquito: enzyme polymorphism and the genetic structure of populations. Ann. Entomol. Soc. Am. 83: 578-585. (1990).

Lewandowski, H.B., Jr., G.R. Hooper, and H.D. Newson. Determination of some important natural potential vectors of dog heartworm in central Michigan. Mosq. News 40: 73-78. (1980).

Maniatis, T., E.F. Fritsch, and J. Sambrook. Molecular cloning: A laboratory manual. Cold Spring Harbor Labs, Cold Spring Harbor, New York. (1982).

Marshall, E. Malaria research -- what next? Science 247: 399-402. (1990).

Mason P.J., and J.O. Bishop. Molecular cloning of part of the mitochondrial DNA of *Drosophila melanogaster.* Biochem. Biophys. Res. Comm. 95: 1268-1274. (1980).

Narang, S.K., P.E. Kaiser, and J.A. Seawright. Dichotomous electrophoretic key for the identification of sibling species A, B and C of the *Anopheles quadrimaculatus* (Say) species complex (Diptera: Culicidae). J. Med. Entomol. 26: 94-99. (1989a).

Narang, S.K., P.E. Kaiser, and J.A. Seawright. Identification of species D, a new member of the *Anopheles quadrimaculatus* species complex: A biochemical key. J. Am. Mosq. Control Assoc. 5: 317-324. (1989b).

Narang, S.K., and J.A. Seawright. Genetic differentiation among members of species complexes in anopheline mosquitoes (Diptera:Culicidae). In: Eukaryotic Structural and Functional Aspects. Narosa-Springers, India. pp. 62-98. (1990).

Nei, M. Molecular Evolutionary Genetics. Columbia University Press. (1987).

Nei, M., and W.-H. Li. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc. Nat. Acad. Sci. USA 76: 5369-5273. (1979).

Nei, M., and F. Tajima. DNA polymorphism detectable by restriction endonucleases. Genetics 97: 145-163. (1981).

Renkawitz, R., and S.A. Gerbi. Ribosomal DNA of the fly *Sciara coprophila* has a very small and homogeneous repeat unit. Mol. Gen. Genet. 173: 1-13. (1979).

Saitou, N., and M. Nei. The neighbor-joining method: a new method for reconstructing phylogenetic trees. J. Mol. Biol. Evol. 4: 406-425. (1987).

Santella, R.M., D. Grunberger, I.B. Weinstein, and A. Rich. Induction of the Z conformation in poly(dG-dC)poly(dG-dC) by binding of N-2-acetylaminofluorene to guanine residues. Proc. Natl. Acad. Sci. USA 78: 1451-1455. (1981).

Saunders, N.C., L.G. Kessler, and J.C. Avise. Genetic variation and geographic differentiation in mitochondrial DNA of the horseshoe crab, *Limulus polyphemus*. Genetics 112: 613-627. (1986).

Shah, D.M., and C.H. Langley. Inter- and intraspecific variation in restriction maps of *Drosophila* mitochondrial DNAs. Nature 281: 696-699. (1979).

Smith, T. F., and M. S. Waterman. Comparison of biosequences. Advances Appl. Math. 2: 482-489. (1981).

Solignac, M., M. Monnerot, and J. Mounolou. Mitochondrial DNA heteroplasmy in *Drosophila mauritiana*. Proc. Natl. Acad. Sci. USA 80: 6942-6946. (1983).

Solignac, M., M. Monnerot, and J. Mounolou. Mitochondrial DNA evolution in the *melanogaster* species subgroup of *Drosophila*. J. Mol. Evol. 23: 31-40. (1986).

Steiner, W.W.M., and D.J. Joslyn. Electrophoretic techniques for the genetic study of mosquitoes. Mosq. News 39: 35-54. (1979).

Wilbur, W. J., and D. J. Lipman. Rapid similarity searches of nucleic acid and protein data banks. Proc. Natl. Acad. Sci. USA 80: 726-730. (1983).

Wilson, A.C., R.L. Cann, S.M. Carr, M. George, U.B. Gyllenstein, K.M. Helmbychowski, R.G. Higuchi, S.R. Palumbi, E.M. Prager, R.D. Sage, and M. Stoneking. Mitochondrial DNA and two perspectives on evolutionary genetics. Biol. J. Linn. Soc. 26: 375-400. (1985).

BIOGRAPHICAL SKETCH

The author was born on September 15, 1951, in Fayetteville, Tennessee. Her family moved to Tallahassee, Florida, in 1957 where she attended elementary and secondary school and received a B.S. in biology from Florida State University in 1975. In 1978 the author moved to Gainesville, Florida, to begin graduate studies at the University of Florida in the Department of Entomology and Nematology. During this time, she was employed by the USDA-ARS, Insects Affecting Man and Animals Research Laboratory, Genetics Research Unit. The author completed her M.S. in entomology in 1984 and continued employment with USDA. In 1987 she began doctoral studies at the University of Florida, Department of Entomology and Nematology. After completing the Ph.D., the author plans to accept a Research Associate position with USDA in population genetics and molecular biology. The author is a member of Gamma Sigma Delta (the honor society of agriculture) and the American Genetic Association.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Jack A. Seawright, Chair
Associate Professor of Entomology
and Nematology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Andrew F. Cockburn, Cochair
Assistant Professor of Entomology
and Nematology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

James E. Maruniak
Associate Professor of
Entomology and Nematology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

H. Glenn Hall
Assistant Professor of
Entomology and Nematology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Christine D. Chase
Assistant Professor of
Horticultural Science

This dissertation was submitted to the Graduate Faculty of the College of Agriculture and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

August 1990

Jack L. Fry

Dean, College of Agriculture

Dean, Graduate School

UNIVERSITY OF FLORIDA
3 1262 08553 8352